

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2001, 14:23:10 ; Search time 87.97 Seconds

(without alignments)
15,624 Million cell updates/sec

Title: PCT-US01-05825A-25

Sequence: 1 EVOLVESGGGLVPGGSLRL 20

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR.67:*

1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1.	98	100.0	97	2	S26890 Ig heavy chain V r
2	98	100.0	97	2	S46462 Ig heavy chain V r
3	98	100.0	98	2	PL0121 Ig heavy chain V-I
4	98	100.0	98	2	PL0123 Ig heavy chain V-I
5	98	100.0	98	2	S26896 Ig heavy chain V r
6	98	100.0	98	2	S29545 Ig heavy chain V r
7	98	100.0	98	2	S26932 Ig heavy chain V r
8	98	100.0	98	2	S26891 Ig heavy chain V r
9	98	100.0	98	2	S26894 Ig heavy chain V r
10	98	100.0	98	2	S26933 Ig heavy chain V r
11	98	100.0	98	2	S26934 Ig heavy chain V r
12	98	100.0	100	2	S26925 Ig heavy chain V r
13	98	100.0	100	2	S26926 Ig heavy chain V r
14	98	100.0	114	2	S36280 Ig heavy chain V r
15	98	100.0	117	2	S17079 Ig heavy chain V-g
16	98	100.0	117	2	A34964 Ig heavy chain pre
17	98	100.0	117	2	S36259 Ig heavy chain V r
18	98	100.0	117	2	S21980 Ig heavy chain V-g
19	98	100.0	118	2	S31121 Ig heavy chain V-h
20	98	100.0	119	1	GHUTE Ig heavy chain V-I
21	98	100.0	119	2	S31107 Ig heavy chain - h
22	98	100.0	120	1	M3HUBW Ig heavy chain V-I
23	98	100.0	121	2	S26798 Ig heavy chain V r
24	98	100.0	123	2	S31509 Ig heavy chain - h
25	98	100.0	123	2	S26794 Ig heavy chain V r
26	98	100.0	124	2	S20775 Ig heavy chain V r
27	98	100.0	125	2	S30531 Ig heavy chain V r
28	98	100.0	136	2	S31587 Ig heavy chain V r
29	98	100.0	137	2	S78054 Ig heavy chain pre

30	98	100.0	139	2	I37781 Ig variable region
31	98	100.0	140	2	S22657 Ig heavy chain pre
32	98	100.0	143	2	S23624 Ig heavy chain V r
33	98	100.0	147	2	I37780 Ig variable region
34	98	100.0	191	2	JL0048 Ig heavy chain V r
35	97	99.0	116	2	B28966 Ig heavy chain pre
36	97	99.0	127	2	PT0369 Ig gamma chain pre
37	95	96.9	82	2	C34964 Ig heavy chain pre
38	95	96.9	97	2	S26886 Ig heavy chain V r
39	95	96.9	98	2	S26889 Ig heavy chain V r
40	95	96.9	98	2	S54856 Ig heavy chain V r
41	95	96.9	100	2	PL0122 Ig heavy chain V-I
42	95	96.9	113	2	S25571 Ig heavy chain V r
43	95	96.9	115	1	H3HUTL Ig heavy chains V-
44	95	96.9	116	1	ALHUTV Ig heavy chain V-I
45	95	96.9	117	1	H3HUT6 Ig heavy chain pre

ALIGNMENTS

RESULT 1
S26890
Ig heavy chain V region (DP-48) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26890
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Dlewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; MUID:93021117
A:Accession: S26890
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-97 <TOM>
A:Cross-References: EMBL:212348; NID:932916; PIDN:CAA78218.1; PID:932917
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 7.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 2
S46462
Ig heavy chain V region (YAC-5) - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C:Accession: S46462
R:COOK, G.P.; Tomlinson, I.M.; Walter, G.; Riethman, H.; Carter, N.P.; Bulwela, L.;
Nature Genet. 7, 162-168, 1994
A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the te
A:Reference number: S46460; MUID:95004581
A:Accession: S46462
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <COO>
A:Cross-References: EMBL:827504; NID:9505430; PIDN:CAA81824.1; PID:9505431
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 7.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGSLRL 20
|||||
Db 1 EVOLVESGGGLVOPGSLRL 20

RESULT 3

PL0121
Ig heavy chain V-III region (TD-Vp) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Aug-1996
C:Accession: PL0121
R:Bird, J.; Gallili, N.; Link, M.; Stiles, D.; Sklar, J.
J. Exp. Med. 168, 229-245, 1988
A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin
A:Reference number: PL0116; MUID:88286083
A:Accession: PL0121
A:Molecule type: mRNA
A:Residues: 1-98 <BIR>
A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:31-35/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGSLRL 20
|||||
Db 1 EVOLVESGGGLVOPGSLRL 20

RESULT 4

PL0123
Ig heavy chain V-III region (TD-Vr) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
C:Accession: PL0123; S26897
R:Bird, J.; Gallili, N.; Link, M.; Stiles, D.; Sklar, J.
J. Exp. Med. 168, 229-245, 1988
A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin
A:Reference number: PL0116; MUID:88286083
A:Accession: PL0123
A:Molecule type: mRNA
A:Residues: 1-98 <BIR>
A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:31-35/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGSLRL 20

Db 1 EVOLVESGGGLVOPGSLRL 20
|||||

RESULT 5

S26896
Ig heavy chain V region (DP-53) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
C:Accession: S26896
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; MUID:93021117
A:Accession: S26896
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12353
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGSLRL 20
|||||
Db 1 EVOLVESGGGLVOPGSLRL 20

RESULT 6

S29545
Ig heavy chain V region (COS 6) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 07-Jan-1994 #sequence_revision 17-Nov-1995 #text_change 23-Jul-1999
C:Accession: S29545
R:Tomlinson, M.; Walter, G.; Cook, G.P.; Winter, G.
Submitted to the EMBL Data Library, October 1992
A:Reference number: S29543
A:Accession: S29545
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z17392; NID:932840; PIDN:CAA78996.1; PID:932841
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGSLRL 20
|||||
Db 1 EVOLVESGGGLVOPGSLRL 20

RESULT 7

S26932
Ig heavy chain V region (DP-39) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26932
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; MUID:93021117
A:Accession: S26932
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:212339; NID:g32898; PIDN:CAA78209.1; PID:g32899
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
|||||
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 8
Ig heavy chain V region (DP-58) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26891
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:212358; NID:g32935; PIDN:CAA78228.1; PID:g32936
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
|||||
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 9
Ig heavy chain V region (DP-51) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26894
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26894
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:212351; NID:g32924; PIDN:CAA78221.1; PID:g32925
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
|||||
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 10
Ig heavy chain V region (DP-40) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26933
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26933
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:212340; NID:g32902; PIDN:CAA78210.1; PID:g32903
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
|||||
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 11
Ig heavy chain V region (DP-59) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26934
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26934
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:212359; NID:g32937; PIDN:CAA78229.1; PID:g32938
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
|||||
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 12
Ig heavy chain V region (DP-29) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26925
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26925
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <TOM>

A:Cross-references: EMBL:Z123331; NID:932879; PIDN:CA478201.1; PID:932880
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 100;
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVOLVESGGGLVOPGSLRL 20
 Db 1 EVOLVESGGGLVOPGSLRL 20

RESULT 13
 S26926
 Ig heavy chain V region (DP-30) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
 C:Accession: S26926
 R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
 J. Mol. Biol. 227, 776-798, 1992
 A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
 A:Reference number: S26885; MUID:93021117
 A:Accession: S26926
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-100 <TOM>
 A:Cross-references: EMBL:Z123332; NID:932883; PIDN:CA478202.1; PID:932884
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:15-100/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 100;
 Best Local Similarity 100.0%; Pred. No. 8.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVOLVESGGGLVOPGSLRL 20
 Db 1 EVOLVESGGGLVOPGSLRL 20

RESULT 14
 S36280
 Ig heavy chain V region (clone alpha-FOG1-A3) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
 C:Accession: S36280
 R:Griffiths, A.D.; Malmqvist, M.; Marks, J.D.; Bye, J.M.; Embleton, M.J.; McCafferty, J.
 EMBO J. 12, 725-734, 1993
 A:Title: Human anti-self antibodies with high specificity from phage display libraries.
 A:Reference number: S36256; MUID:93178448
 A:Accession: S36280
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-114 <GR1>
 A:Cross-references: EMBL:Z18822
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:15-98/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 114;
 Best Local Similarity 100.0%; Pred. No. 9.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVOLVESGGGLVOPGSLRL 20
 Db 1 EVOLVESGGGLVOPGSLRL 20

RESULT 15
 S17079
 Ig heavy chain V-gene (clone HHG19) - human
 C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
 C:Accession: S60299; S17079
 R:Kueppers, R.; Fischer, U.; Rajewsky, K.; Gause, A.
 Immunol. Lett. 34, 57-62, 1992
 A:Title: Immunoglobulin heavy and light chain gene sequences of a human CD5 positive
 A:Reference number: S60295; MUID:93122853
 A:Accession: S60299
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-117 <KU2>
 A:Cross-references: EMBL:X62128; NID:938340; PIDN:CA444059.1; PID:938341
 A:Note: The authors did not translate the codons for residues 6, 52, 54, 68, 69, 71.
 C:Genetics: 16/1
 A:Introns: 16/1
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 98; DB 2; Length 117;
 Best Local Similarity 100.0%; Pred. No. 9.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVOLVESGGGLVOPGSLRL 20
 Db 20 EVOLVESGGGLVOPGSLRL 39

Search completed: June 13, 2001, 14:23:11
 Job time: 742 sec

This Page Blank (uspto)

RT "Comparative studies on monocytic IgM lambda and IgG kappa from an
RT individual patient. III. The complete amino acid sequence of the VH
RT region of the IgM paraprotein.";
RL Immunochimistry 13:995-999(1976).
CC -I- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE
CC SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM
CC TYPE.
CC PIR: A02049; M3HDM.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 120
SQ SEQUENCE 120 AA; 13227 MW; D3F0428F7C2E6410 CRC64;

Query Match 100.0%; Score 98; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 5.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 3
HV05_CARAU STANDARD; PRT; 116 AA.
ID HV05_CARAU
AC P19181;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 5A PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8814476; PubMed=3125551;
RA Wilson M.R., Middleton D., Marr G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure
RT and family relationships of two genes and a pseudogene in a teleost
RT fish.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
DR PIR: B28966; B28966.
DR InterPro: IPR003006; -
KW Immunoglobulin V region; Signal.
FT CHAIN 1
FT DOMAIN 20 116 IG HEAVY CHAIN V REGION 5A.
FT DOMAIN 20 49 FRAMEWORK 1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 55 68 FRAMEWORK 2.
FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 85 116 FRAMEWORK 3.
FT DISULFID 41 114 BY SIMILARITY.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;

Query Match 99.0%; Score 97; DB 1; Length 116;
Best Local Similarity 95.0%; Pred. No. 7.4e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
DB 20 EVOLVESGGGLVPGGSLRL 39

RESULT 4
HV3D_HUMAN STANDARD; PRT; 115 AA.
ID HV3D_HUMAN

AC P01765;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=78005528; PubMed=409716;
RA Wang A.-C., Wang I.Y., Fudenberg H.H.;
RT "Immunoglobulin structure and genetics. Identity between variable
RT regions of a mu and a gamma2 chain.";
RL J. Biol. Chem. 252:7192-7199(1977).
CC -I- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
CC OF IGM AND IGC2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
CC GAMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
CC IDENTICAL.
DR PIR: A02048; H3HUTL.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 115
SQ SEQUENCE 115 AA; 12356 MW; 4DC67D179F62326 CRC64;

Query Match 96.9%; Score 95; DB 1; Length 115;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 5
HV3R_HUMAN STANDARD; PRT; 116 AA.
ID HV3R_HUMAN
AC P1779;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kehoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains
RT of the VH3 subgroup: definitive identification of four heavy chain
RT hypervariable regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGAL MYELOMA
CC PROTEIN.
DR PIR: A02062; A1H0TV.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12431 MW; EB705F553A963F0C CRC64;

Query Match 96.9%; Score 95; DB 1; Length 116;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVOLVESGGGLVPGGSLRL 20
DB 1 EVOLVESGGGLVPGGSLRL 20

```

DB      1 EVOLLESGGGLVOPGSLRL 20

RESULT  6
HV3C_HUMAN  STANDARD:      PRT:      117 AA.
AC      P01764;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      IG HEAVY CHAIN V-III REGION VH26 PRECURSOR.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=81101090; PubMed=6450418;
RA      Mathysens G., Rabbits T.H.;
RT      "Structure and multiplicity of genes for the human immunoglobulin
RT      heavy chain variable region."
RL      Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: J00236; AAA53516.1; -
DR      EMBL: M35415; AAA58735.1; -
DR      PIR: A02047; H3H26.
DR      InterPro: IPR003006; -.
KW      Immunoglobulin V region; Signal.
FT      SIGNAL 1 19
FT      CHAIN 20 117 IG HEAVY CHAIN V-III REGION VH26.
FT      NON_TER 117 117
SQ      SEQUENCE 117 AA; 12582 MW; E82673F1A3CB0F1 CRC64;

Query Match          96.9%; Score 95; DB 1; Length 117;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 EVOLVESGGGLVOPGSLRL 20
DB      20 EVOLLESGGGLVOPGSLRL 39

RESULT  7
HV3O_HUMAN  STANDARD:      PRT:      117 AA.
AC      P01776;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      IG HEAVY CHAIN V-III REGION WAS.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE.
RX      MEDLINE=74142702; PubMed=4522793;
RA      Capra J.D., Kehoe J.M.;
RT      "Variable region sequences of five human immunoglobulin heavy chains
RT      of the VH3 subgroup: definitive identification of four heavy chain
RT      hypervariable regions."
RL      Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC      -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA

```

```

CC      PROTEIN.
DR      PIR: A02058; G1H0WS.
DR      HSSP: P01772; 2IG2.
DR      InterPro: IPR003006; -.
DR      Pfam: PF00047; 1g; 1.
KW      Immunoglobulin V region.
FT      NON_TER 117 117
SQ      SEQUENCE 117 AA; 13091 MW; 201DF0E1E53D9BF CRC64;

Query Match          96.9%; Score 95; DB 1; Length 117;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 EVOLVESGGGLVOPGSLRL 20
DB      1 EVOLLESGGGLVOPGSLRL 20

RESULT  8
HV3M_HUMAN  STANDARD:      PRT:      119 AA.
AC      P01774;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      IG HEAVY CHAIN V-III REGION POM.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE.
RX      MEDLINE=75046755; PubMed=4139708;
RA      Capra J.D., Kehoe J.M.;
RT      "Structure of antibodies with shared idiotypy: the complete sequence
RT      of the heavy chain variable regions of two immunoglobulin M
RT      anti-gamma globulins."
RL      Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).
CC      -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
CC      GLOBULIN ACTIVITY.
DR      PIR: A02057; M3HUPM.
DR      HSSP: P01772; 2IG2.
DR      InterPro: IPR003006; -.
DR      Pfam: PF00047; 1g; 1.
KW      Immunoglobulin V region.
FT      VARIANT 54 54
FT      NON_TER 119 119
FT      SEQUENCE 119 AA; 12953 MW; 2E018AF4DCEB2610 CRC64;

Query Match          96.9%; Score 95; DB 1; Length 119;
Best Local Similarity 95.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 EVOLVESGGGLVOPGSLRL 20
DB      1 EVOLLESGGGLVOPGSLRL 20

RESULT  9
HV3F_HUMAN  STANDARD:      PRT:      115 AA.
AC      P01767;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      IG HEAVY CHAIN V-III REGION BDY.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78137069; PubMed=416441;
 RA Toriano A., Putnam F.W.;
 RT "Complete amino acid sequence of the alpha 2 heavy chain of a human
 RL 1A2 immunoglobulin of the A2m (2) allotype.";
 CC Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
 CC -1- MISCELLANEOUS: THE SEQUENCE OF THE ALPHA-2, A2M(2) ALLOTYPE, C
 DR REGION OF THIS MYELOMA PROTEIN IS ALSO GIVEN.
 PIR: A02050; A2HBU
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT NON_TER 115
 SQ SEQUENCE 115 AA; 12379 MW; 208876A7DF52DC64 CRC64;

Query Match 95.9%; Score 94; DB 1; Length 115;
 Best Local Similarity 90.0%; Pred. No. 2.1e-07;
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVQLVSGGGLVOPGGSRL 20
 DB 1 EVQLVETGGGLIOPGGSRL 20

RESULT 10
 HV17_MOUSE STANDARD; PRT; 117 AA.
 AC P01786;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V REGION MOPC 47A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=80049769; PubMed=115869;
 RA Robinson E.A., Appella E.;
 RT "Amino acid sequence of a mouse myeloma immunoglobulin heavy chain
 RL (MOPC 47 A) with a 100-residue deletion.";
 RL J. Biol. Chem. 254:11418-11430(1979).
 CC -1- MISCELLANEOUS: THIS ALPHA CHAIN WAS ISOLATED FROM A MYELOMA
 CC PROTEIN THAT CONTAINS ONE LIGHT AND ONE HEAVY CHAIN PER MOLECULE,
 CC LINKED BY A DISULFIDE BOND. IN CONTRAST, NORMAL MOUSE IGA
 CC MOLECULES CONTAIN TWO LIGHT AND TWO HEAVY CHAINS AND LACK A
 CC LIGHT-HEAVY CHAIN DISULFIDE BOND.
 PIR: A02069; A1MS47.
 DR HSSP; P01789; 2MCP.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region.
 FT NON_TER 117
 SQ SEQUENCE 117 AA; 12975 MW; 0C74BE8B154BDF4 CRC64;

Query Match 95.9%; Score 94; DB 1; Length 117;
 Best Local Similarity 95.0%; Pred. No. 2.1e-07;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVQLVSGGGLVOPGGSRL 20
 DB 1 EVKLVEGGGLVOPGGSRL 20

RESULT 11
 HV20_MOUSE STANDARD; PRT; 122 AA.
 AC P01789;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V REGION M603.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=80199926; PubMed=6769593;
 RA Early P., Huang H., Davis M., Calame K., Hood L.;
 RT "An immunoglobulin heavy chain variable region gene is generated from
 RL three segments of DNA: VH, D and JH.";
 RL Cell 19:981-992(1980).
 CC [2]
 CC SEQUENCE OF 1-120.
 RX MEDLINE=75017346; PubMed=4213527;
 RA Rudikoff S., Potter M.;
 RT "Variable region sequence of the heavy chain from a phosphorylcholine
 RL binding myeloma protein.";
 RL Biochemistry 13:4033-4036(1974).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF FAB FRAGMENT.
 RX MEDLINE=75065510; PubMed=4530984;
 RA Segal D.M., Padlan E.A., Cohen G.H., Rudikoff S., Potter M.,
 RA Davies D.R.;
 RT "The three-dimensional structure of a phosphorylcholine-binding mouse
 RL immunoglobulin Fab and the nature of the antigen binding site.";
 CC Proc. Natl. Acad. Sci. U.S.A. 71:4298-4302(1974).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
 CC BINDS PHOSPHORYLCHOLINE.
 PIR: A02070; AVMST5.
 DR PDB; 1MCP; 15-JUL-92.
 DR PDB; 2MCP; 15-JUL-92.
 DR InterPro: IPR003006;
 DR Pfam: PF00047; 19; 1.
 KW Immunoglobulin V region; 3D-structure.
 FT SITE 33
 FT SITE 33
 FT SITE 52
 FT SITE 52
 FT STRAND 3
 FT STRAND 7
 FT TURN 11
 FT TURN 12
 FT TURN 14
 FT STRAND 15
 FT STRAND 17
 FT HELIX 25
 FT STRAND 29
 FT STRAND 31
 FT TURN 33
 FT TURN 39
 FT TURN 41
 FT STRAND 42
 FT TURN 46
 FT TURN 50
 FT STRAND 55
 FT TURN 61
 FT STRAND 66
 FT STRAND 67
 FT TURN 68
 FT STRAND 69
 FT STRAND 70
 FT TURN 75
 FT STRAND 76
 FT TURN 79
 FT STRAND 80
 FT HELIX 86
 FT STRAND 90
 FT STRAND 92
 FT STRAND 94
 FT STRAND 103
 FT STRAND 107
 FT STRAND 112
 FT STRAND 116
 FT NON_TER 122
 SQ SEQUENCE 122 AA; 13626 MW; BA2C864438B64F0F CRC64;

Query Match 95.9%; Score 94; DB 1; Length 122;
 Best Local Similarity 95.0%; Pred. No. 2.2e-07;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVQLVSGGGLVOPGGSRL 20
 DB 1 EVKLVEGGGLVOPGGSRL 20

```

RESULT 12
HV1_MOUSE      STANDARD;      PRT;      122 AA.
ID HV21_MOUSE
AC P01790;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION M511.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=81054880; PubMed=6776528;
RA Robinson E.A., Appella E.;
RT "Complete amino acid sequence of a mouse immunoglobulin alpha chain
  (MOPC 511).";
RL Proc. Natl. Acad. Sci. U.S.A. 77:4909-4913(1980).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
  BINDS PHOSPHORYLCHOLINE.
DR PIR: A02070; AVMS75.
DR HSSP: P01789; 2MCP.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region.
FT NON_TER 122
SQ SEQUENCE 122 AA; 13652 MW; 9FA837731EA50207 CRC64;

```

```

Query Match      95.9%; Score 94; DB 1; Length 122;
Best Local Similarity 95.0%; Pred. No. 2.2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVLVESGGGLVPGGSLRL 20
Db 1 EVLVESGGGLVPGGSLRL 20

```

```

RESULT 13
HV18_MOUSE      STANDARD;      PRT;      123 AA.
ID HV18_MOUSE
AC P01787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGIONS TERC 15/S107/HPCML/HPCM2/HPCM3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE (TERC 15).
RX MEDLINE=76222762; PubMed=8199332;
RA Rudnikoff S., Potter M.;
RT "Size differences among immunoglobulin heavy chains from
  phosphotyrosine-binding proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 73:2109-2112(1976).
RN [2]
RP SEQUENCE FROM N.A. (H107).
RX MEDLINE=80199926; PubMed=6769593;
RA Early P., Huang H., Davis M., Calame K., Hood L.;
RT "An immunoglobulin heavy chain variable region gene is generated from
  three segments of DNA: VH, D and JH.";
RL Cell 19:981-992(1980).
RN [3]
RP SEQUENCE (S107).
RX MEDLINE=76110488; PubMed=813561;
RA Rudnikoff S., Barstad P., Potter M., Hood L.;
RL Unpublished results, cited by:
  Hood L., Campbell J.H., Elgin S.C.R.;

```

```

RL Annu. Rev. Genet. 9:305-353(1975).
RN [4]
RP SEQUENCE (HPCML; HPCM2 AND HPCM3).
RX MEDLINE=8197602; PubMed=7231520;
RA Gearhart P.J., Johnson N.D., Douglas R., Hood L.;
RT "Igg antibodies to phosphorylcholine exhibit more diversity than
  their Igm counterparts.";
RL Nature 291:29-34(1981).
CC -1- MISCELLANEOUS: ALL THOSE SEQUENCE APPEARS TO BE IDENTICAL.
CC -1- MISCELLANEOUS: THESE CHAINS WERE ISOLATED FROM MYELOMA AND
  HYBRIDOMA PROTEINS THAT BIND PHOSPHORYLCHOLINE.
DR PIR: A02070; AVMS75.
DR HSSP: P01789; 2MCP.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region; Hybridoma.
FT NON_TER 123
SQ SEQUENCE 123 AA; 13777 MW; 9D58086DE12F7000 CRC64;

```

```

Query Match      95.9%; Score 94; DB 1; Length 123;
Best Local Similarity 95.0%; Pred. No. 2.2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVLVESGGGLVPGGSLRL 20
Db 1 EVLVESGGGLVPGGSLRL 20

```

```

RESULT 14
HV19_MOUSE      STANDARD;      PRT;      123 AA.
ID HV19_MOUSE
AC P01788;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION H8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RA Barstad P.;
RL Thesis (1975), California Institute of Technology / Pasadena, U.S.A.
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
  BINDS PHOSPHORYLCHOLINE.
DR PIR: A02070; AVMS75.
DR HSSP: P01789; 2MCP.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 1.
KW Immunoglobulin V region.
FT NON_TER 123
SQ SEQUENCE 123 AA; 13805 MW; 9D581401912F7000 CRC64;

```

```

Query Match      95.9%; Score 94; DB 1; Length 123;
Best Local Similarity 95.0%; Pred. No. 2.2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVLVESGGGLVPGGSLRL 20
Db 1 EVLVESGGGLVPGGSLRL 20

```

```

RESULT 15
HV22_MOUSE      STANDARD;      PRT;      123 AA.
ID HV22_MOUSE
AC P01791;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION HPCM6.

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE-81197602; PubMed-7231520;
RA Gearhart P.J., Johnson N.D., Douglas R., Hood L.;
RT "IgG antibodies to phosphorylcholine exhibit more diversity than
RL their IgM counterparts.";
RL Nature 291:29-34(1981)
CC -! MISCELANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS PHOSPHORYLCHOLINE.
DR PIR: A02070; AVNSTS.
DR HSSP; P01789; 2MCP.
DR InterPro; IPR03006; -.
DR Pfam; PF00047; 1g; 1.
KW Immunoglobulin V region; Hybridoma.
FT NON_TER 123
SO SEQUENCE 123 AA; 13895 MW; 81361892CBF7000 CRC64;

Query Match 95.9%; Score 94; DB 1; Length 123;
Best Local Similarity 95.0%; Pred. No. 2.2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVOPGGSRL 20
||:|||||
Db 1 EVKLVESGGGLVOPGGSRL 20

Search completed: June 13, 2001, 14:30:36
Job time: 526 sec

DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035042; AAD56278.1; -
 DR HSSP: P01772; 2F84.
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 118 118
 SQ SEQUENCE 118 AA; 12872 MW; BAD1A5944BD5CCA CRC64;

Query Match
 Best Local Similarity 100.0%; Score 98; DB 4; Length 118;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVOLVESGGGLVPGGSLRL 20
 DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 3
 Q9UL90 PRELIMINARY; PRT; 113 AA.
 AC Q9UL90;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035024; AAD56260.1; -
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA; 12437 MW; ED57EDD19086D07F CRC64;

Query Match
 Best Local Similarity 96.9%; Score 95; DB 4; Length 113;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVOLVESGGGLVPGGSLRL 20
 DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 4
 Q9UL71 PRELIMINARY; PRT; 121 AA.
 ID Q9UL71

AC Q9UL71;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035043; AAD56279.1; -
 DR HSSP: P01772; 2F84.
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 121 121
 SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFA5D50736 CRC64;

Query Match
 Best Local Similarity 96.9%; Score 95; DB 4; Length 121;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVOLVESGGGLVPGGSLRL 20
 DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 5
 Q9UL88 PRELIMINARY; PRT; 131 AA.
 ID Q9UL88;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035026; AAD56262.1; -
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 14142 MW; 96E7D668E375DEAO CRC64;

Query Match
 Best Local Similarity 95.0%; Score 94; DB 4; Length 131;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVOLVESGGGLVPGGSLRL 20
 DB 1 EVOLVESGGGLVPGGSLRL 20

RESULT 6

09ULB6
ID 09ULB6 PRELIMINARY: PRT: 95 AA.
AC 09ULB6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE IMNOGLOBULIN HEAVY CHAIN (FRAGMENT).
GN VH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tange Y., Kayano H.;
RT "Human VH gene sequence."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB035268; BAA87067.1; -.
DR HSSP: P01772; 2F84
DR INTERPRO: IPR003006; -.
DR PFM: PF00047; 19; 1.
FT NON_TER 1
FT NON_TER 95
SQ SEQUENCE 95 AA: 10527 MW: 90A8C6D16D22574A CRC64;

Query Match 88.8%; Score 87; DB 4; Length 95;
Best Local Similarity 94.7%; Pred. No. 1.4e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 VOLVESGGGLVOPGSLRL 20
DB 1 VOLVESGGGLVOPGSLRL 19
|||||

RESULT 7

ID 09ULB4 PRELIMINARY: PRT: 122 AA.
AC 09ULB4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035030; AAD56266.1; -.
DR HSSP: P01772; 2F84
DR INTERPRO: IPR003006; -.
DR PFM: PF00047; 19; 1.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA: 13579 MW: 36054D41366545B8 CRC64;

Query Match 88.8%; Score 87; DB 4; Length 122;
Best Local Similarity 90.0%; Pred. No. 1.8e-05;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGSLRL 20
DB 1 EVOLVESGGGLVOPGSLRL 20
|||||

RESULT 8
ID 09UL93 PRELIMINARY: PRT: 116 AA.
AC 09UL93;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035021; AAD56257.1; -.
DR INTERPRO: IPR003006; -.
DR PFM: PF00047; 19; 1.
FT NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA: 12434 MW: 0DA0348154DB6061 CRC64;

Query Match 83.7%; Score 82; DB 4; Length 116;
Best Local Similarity 89.5%; Pred. No. 9e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VOLVESGGGLVOPGSLRL 20
DB 1 VOLVESGGGLVOPGSLRL 19
|||||

RESULT 9

ID 09RI44 PRELIMINARY: PRT: 437 AA.
AC 09RI44;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GAMMA1 HEAVY CHAIN OF MAb7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv)."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF152372; AAD40243.1; -.
DR HSSP: P01842; 7FAB
DR INTERPRO: IPR003006; -.
DR PFM: PF00047; 19; 4.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA: 48142 MW: 5C3A7BB3EE7D697C CRC64;

Query Match 81.6%; Score 80; DB 11; Length 437;
Best Local Similarity 84.2%; Pred. No. 0.00076;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VOLVESGGGLVOPGSLRL 20
DB 1 VOLVESGGGLVOPGSLRL 19
|||||

```

RESULT 10
Q9Y509 PRELIMINARY: PRT: 147 AA.
AC Q9Y509:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VH3 PROTEIN (FRAGMENT).
GN VH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96071149; PubMed=7475288;
RA Cao J., Vescio R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C.,
RA Lichtenstein A.K., Berenson J.R.;
RT "A CD10-positive subset of malignant cells is identified in multiple
RT myeloma using PCR with patient-specific immunoglobulin gene primers.";
RL Leukemia 9:1948-1953(1995).
DR EMBL: S80860; AAD14339.1; -.
DR HSSP: P01772; 2FB4.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 1.
FT NON_TER
SQ SEQUENCE 147 AA; 15768 MW; 8489FCAA7BC925C CRC64;

```

```

Query Match
Best Local Similarity 80.6%; Score 79; DB 4; Length 147;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVPGGSLRL 20
Db 1 QVHLVSGGGLVPGGSLRL 20

```

```

RESULT 11
Q9OV16 PRELIMINARY: PRT: 15 AA.
AC Q9OV16:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PROLACTIN-BINDING PROTEIN (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=95094032; PubMed=8000909;
RA Cohen H., Cohen O., Gagnon J.;
RT "Serum prolactin-binding protein (PRL-BP) of human and rat are
RT identified as IgG.";
RL C. R. Acad. Sci., III, Sci. Vie 317:293-298(1994).
DR HSSP: P01789; IMCP.
SQ SEQUENCE 15 AA; 1469 MW; 35ED2512FF3FA369 CRC64;

```

```

Query Match
Best Local Similarity 76.5%; Score 75; DB 11; Length 15;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVPG 15
Db 1 EVOLVESGGGLVPG 15

```

RESULT 12

```

Q9QYF0 PRELIMINARY: PRT: 298 AA.
AC Q9QYF0:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CN 8 SCFV.
GN CN 8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=SPLEEN;
RA Shinohara N., Demura T., Fukuda H.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=SPLEEN;
RA Shinohara N., Demura T., Fukuda H.;
RT "Isolation of a novel type of vascular cell wall-specific monoclonal
RT antibody recognizing a cell polarity using a phage display subtraction
RT method.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB036341; BA88633.1; -.
DR HSSP: P01607; IREI.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 2.
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

```

```

Query Match
Best Local Similarity 76.5%; Score 75; DB 11; Length 298;
Matches 14; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVPGGSLRL 20
Db 40 QVRLQSGGGLVPGGSLRL 59

```

```

RESULT 13
Q9N0M4 PRELIMINARY: PRT: 124 AA.
AC Q9N0M4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ANTI-HUMAN A33 HEAVY CHAIN DOMAIN (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Rader C., Rittler G., Nathan S., Elia M., Gout I., Junghuth A.A.,
RA Cohen L.S., Welt S., Old L.J., Barbas C.F., III.;
RT "The rabbit antibody repertoire as a novel source for the generation
RT of therapeutic human antibodies.";
RL J. Biol. Chem. 275:13668-13676(2000).
DR EMBL: AF245503; AAF68450.1; -.
FT NON_TER
SQ SEQUENCE 124 AA; 13476 MW; 96D2B29FE27C24C8 CRC64;

```

```

Query Match
Best Local Similarity 71.4%; Score 70; DB 6; Length 124;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 QLVESGGGLVPGGSLRL 20
Db 3 QLVESGGGLVPGGSLRL 20

```

RESULT 14

Q9N0W6 PRELIMINARY; PRT; 124 AA.
 AC Q9N0W6:
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE ANTI-HUMAN A33 HEAVY CHAIN DOMAIN (FRAGMENT).
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rader C., Rilter G., Nathan S., Elia M., Gout I., Jungbluth A.A.,
 RA Cohen L.S., Wolt S., Old L.J., Barbas C.F. III.;
 RT "Type rabbit antibody repertoire as a novel source for the generation
 RT of therapeutic human antibodies."
 RL J. Biol. Chem. 275:13668-13676(2000).
 DR EMBL; AF245501; AAF68448.1; -.
 FT NON_TER 1
 FT 124
 SO SEQUENCE 124 AA; 13508 MW; 96C9B29FE26724C3 CRC64;

Query Match

Best Local Similarity 68.4%; Score 67; DB 6; Length 124;

Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 OLVESSGGGLVQPGSLRL 20

DB 3 QLMESGGGLVTLGGSLKL 20

RESULT 15

Q9UC53

PRELIMINARY; PRT; 16 AA.

AC Q9UC53:
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE 77 KDA SPONTANEOUS RECURRENT ABORTION-ASSOCIATED HUMAN EMBRYONIC
 DE ANTIGEN/IGWHIT1 HOMOLOG (FRAGMENT).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-96033130; PubMed-8582963;
 RA Shiraishi Y., Shiraishi Y., Yamamoto D., Hasegawa T., Kitamura W.,
 RA Miki S., Tanaka T., Suzuki T., Soma H.;
 RT "Diagnostic relevance of abortion-associated human embryonic antigen
 RT expressed on the cell surface of tumour promoter-treated Bloom
 RT syndrome cells."
 RL Hum. Reprod. 10:1694-1701(1995).
 SO SEQUENCE 16 AA; 1626 MW; C9C5ED2512FF3FB9 CRC64;

Query Match

Best Local Similarity 67.3%; Score 66; DB 4; Length 16;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVQPG 15

DB 1 EVOLVESGGGLVQPG 15

Search completed: June 13, 2001, 14:29:43
 Job time: 544 sec

RESULT 2
US-08-471-780C-80
; Sequence 80, Application US/08471780C
; Patent No. 5759808
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,780C
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/106,944
; FILING DATE: 17-AUG-1993
; APPLICATION NUMBER: FR 92402326.0
; FILING DATE: 21-AUG-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: FR 93401310.3
; FILING DATE: 21-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 04958.0008-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Camelus dromedarius
; US-08-471-780C-80

Query Match 100.0%; Score 98; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVQPGGSLRL 20
DB 1 EVOLVESGGGLVQPGGSLRL 20

RESULT 3
US-08-467-282B-80
; Sequence 80, Application US/08467282B
; Patent No. 5800988
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,282B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/106,944
; FILING DATE: 17-AUG-1993
; APPLICATION NUMBER: FR 92402326.0
; FILING DATE: 21-AUG-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: FR 93401310.3
; FILING DATE: 21-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 04958.0008-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Camelus dromedarius
; US-08-467-282B-80

Query Match 100.0%; Score 98; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVQPGGSLRL 20
DB 1 EVOLVESGGGLVQPGGSLRL 20

RESULT 4
US-08-471-282A-80
; Sequence 80, Application US/08471282A
; Patent No. 5840853
; GENERAL INFORMATION:
; APPLICANT: Casterman, Cecile
; APPLICANT: Hamers, Raymond
; TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,282A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-282A-80

Query Match 100.0%; Score 98; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 3,8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EVOLVESGGGLVPGGSLRL 20
|||||
Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 5
US-08-466-710C-80
Sequence 80, Application US/08466710C
Patent No. 5874541
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pinegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,710C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-466-710C-80

Query Match 100.0%; Score 98; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 3,8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EVOLVESGGGLVPGGSLRL 20
|||||
Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 6
US-08-468-739C-80
Sequence 80, Application US/08468739C
Patent No. 6015695
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pinegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,739C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-468-739C-80

Query Match 100.0%; Score 98; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGGSRL 20
|||||
DB 1 EVOLVESGGGLVOPGGSRL 20

RESULT 7
US-08-470-139-17
Sequence 17, Application US/08470139
Patent No. 5998586
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Interleukin-5 specific recombinant antibodies
NUMBER OF SEQUENCES: 28
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,139
FILING DATE: 06 JUNE-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: TRUJILLO, DOREEN YAKKO
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0044
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-470-139-17

Query Match 100.0%; Score 98; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGGSRL 20
|||||
DB 1 EVOLVESGGGLVOPGGSRL 20

RESULT 8
US-08-211-202-116
Sequence 116, Application US/08211202
Patent No. 5565332
GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
APPLICANT: BAIRER, Michael
APPLICANT: JESPER, Laurent Stephane Anne Therese
APPLICANT: WINTER, Gregory Paul
TITLE OF INVENTION: Production of chimeric antibodies - a
TITLE OF INVENTION: combinatorial approach
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
ADDRESS: Botun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois

COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120252.3
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120377.8
FILING DATE: 25-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-211-202-116

Query Match 100.0%; Score 98; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGGGLVOPGGSRL 20
|||||
DB 1 EVOLVESGGGLVOPGGSRL 20

RESULT 9
US-07-942-245-37
Sequence 37, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Brydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughree Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-37

Query Match 100.0%; Score 98; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
|||||
Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 10
US-08-665-202-31
Sequence 31, Application US/08665202

PATENT NO. 5977322
GENERAL INFORMATION:
APPLICANT: Marks, James D.
TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
TITLE OF INVENTION: Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,202
FILING DATE: 13-JUN-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061410
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-665-202-31

Query Match 100.0%; Score 98; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
|||||
Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 11
US-08-428-197-4
Sequence 4, Application US/08428197

PATENT NO. 5891438
GENERAL INFORMATION:
APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,197
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
CLONE: SFL
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..108
US-08-428-197-4

Query Match 100.0%; Score 98; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20
|||||

Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 12

PCT-US93-10555-4

Sequence 4, Application PC/TUS9310555

GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.

TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF

TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH

TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Spensley Horn Judas & Lubitz

STREET: 1880 Century Park East - Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10555

FILING DATE: 29-OCT-1993

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Howells, Stacy L.

REGISTRATION NUMBER: 34,842

REFERENCE/DOCKET NUMBER: FD-2630

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 108 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

CLONE: SFL

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..108

PCT-US93-10555-4

Query Match 100.0%; Score 98; DB 5; Length 108;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20

Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 13

US-08-428-197-3

Sequence 3, Application US/08428197

Patent No. 5891438

GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.

TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF

TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH

TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Spensley Horn Judas & Lubitz

STREET: 1880 Century Park East - Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10555

FILING DATE: 29-OCT-1993

Prior Application DATA:

APPLICATION NUMBER: PCT/US93/10555

FILING DATE: 29-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Howells, Stacy L.

REGISTRATION NUMBER: 34,842

REFERENCE/DOCKET NUMBER: FD-2630

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 109 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE:

IMMEDIATE SOURCE:

CLONE: Huab14-3

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..109

US-08-428-197-3

Query Match 100.0%; Score 98; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGGGLVPGGSLRL 20

Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 14

PCT-US93-10555-3

Sequence 3, Application PC/TUS9310555

GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.

TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF

TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH

TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Spensley Horn Judas & Lubitz

STREET: 1880 Century Park East - Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10555

FILING DATE: 29-OCT-1993

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: Huab14-3
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..109
PCT-US93-10555-3

Query Match 100.0%; Score 98; DB 5; Length 109;
Best Local Similarity 100.0%; Pred.No. 1.7e-07;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVPGGSLRL 20
|||
Db 1 EVOLVESGGGLVPGGSLRL 20

RESULT 15
US-08-974-899-6
Sequence 6, Application US/08974899
Patent No. 6037454
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Jardieu, Paula M.
TITLE OF INVENTION: Humanized Anti-CD11a Antibodies
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,899
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/031971
FILING DATE: 11/27/96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1014R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1994
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-974-899-6

Query Match 100.0%; Score 98; DB 3; Length 113;
Best Local Similarity 100.0%; Pred.No. 1.8e-07;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVOLVESGGGLVPGGSLRL 20
|||
Db 1 EVOLVESGGGLVPGGSLRL 20

Search completed: June 13, 2001, 14:27:07
Job time: 628 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:45 ; Search time 150.28 Seconds
(without alignments)
3.423 Million cell updates/sec

Title: PCT-US01-05825A-26

Perfect score: 45

Sequence: 1 VTFEYDPAVS 9

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseqp/AA1987.DAT.*
8: /SID56/gcgdata/geneseq/geneseqp/AA1988.DAT.*
9: /SID56/gcgdata/geneseq/geneseqp/AA1989.DAT.*
10: /SID56/gcgdata/geneseq/geneseqp/AA1990.DAT.*
11: /SID56/gcgdata/geneseq/geneseqp/AA1991.DAT.*
12: /SID56/gcgdata/geneseq/geneseqp/AA1992.DAT.*
13: /SID56/gcgdata/geneseq/geneseqp/AA1993.DAT.*
14: /SID56/gcgdata/geneseq/geneseqp/AA1994.DAT.*
15: /SID56/gcgdata/geneseq/geneseqp/AA1995.DAT.*
16: /SID56/gcgdata/geneseq/geneseqp/AA1996.DAT.*
17: /SID56/gcgdata/geneseq/geneseqp/AA1997.DAT.*
18: /SID56/gcgdata/geneseq/geneseqp/AA1998.DAT.*
19: /SID56/gcgdata/geneseq/geneseqp/AA1999.DAT.*
20: /SID56/gcgdata/geneseq/geneseqp/AA2000.DAT.*
21: /SID56/gcgdata/geneseq/geneseqp/AA2001.DAT.*
22: /SID56/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	20 W94488	Human adult brain
2	45	100.0	9	21 W79131	Human adult brain
3	34	75.6	205	20 W41173	Llama Vhh polypept
4	34	75.6	1475	11 R08221	Recombinant alpha
5	32	71.1	183	20 V41166	Llama Vhh polypept
6	32	71.1	477	15 R47450	T84.12 Heavy chain
7	32	71.1	477	15 R47453	ChIT84.12 H3 heavy
8	31	68.9	119	18 W13514	Anti-melanoma anti
9	31	68.9	208	20 V41168	Llama Vhh polypept
10	31	68.9	457	20 V49069	Amino acid sequenc
11	31	68.9	466	20 V31769	Staphylococcus aur

12	31	68.9	466	21 Y70133	Staphylococcus aur
13	30	66.7	37	21 B33998	Human secreted pro
14	30	66.7	48	20 Y41320	Human secreted pro
15	30	66.7	56	20 Y02500	Clone selected aft
16	30	66.7	113	12 R15677	SYNTH-9 Bt2 modif
17	30	66.7	116	20 Y06128	Monoclonal antibod
18	30	66.7	190	20 Y37292	Protein which is s
19	30	66.7	289	20 Y06132	Monoclonal antibod
20	30	66.7	365	19 W98527	H. pylori GHP0 140
21	30	66.7	436	19 W49680	Open reading frame
22	30	66.7	452	20 Y09553	Streptococcus pyog
23	30	66.7	469	21 Y95547	Chlamydia pneumoni
24	30	66.7	593	17 R97661	Bacillus thuringie
25	30	66.7	594	14 R37267	Truncated modified
26	30	66.7	594	15 R63209	CYIA(b)(5.3) toxi
27	30	66.7	595	15 R63210	CYIA(c)(6.6) toxi
28	30	66.7	595	17 R97662	Bacillus thuringie
29	30	66.7	610	7 P60094	Sequence of the Ba
30	30	66.7	612	21 B13018	HD-73 delta-endoto
31	30	66.7	613	16 R65296	Bacillus thuringie
32	30	66.7	613	21 Y76917	Synthetic Bacillus
33	30	66.7	613	21 Y54826	Synthetic Bt endot
34	30	66.7	614	16 R65297	Bacillus thuringie
35	30	66.7	614	21 Y76918	Synthetic Bacillus
36	30	66.7	614	21 Y54828	Synthetic Bt endot
37	30	66.7	615	21 B18416	Amino acid sequenc
38	30	66.7	617	18 W14699	CYIA(c) crystal p
39	30	66.7	622	14 R37266	Truncated insectic
40	30	66.7	650	21 Y56096	B. thuringiensis C
41	30	66.7	724	15 R46232	Bacillus thuringie
42	30	66.7	869	10 P93171	Half-length hybrid
43	30	66.7	869	21 Y90705	Half-length hybrid
44	30	66.7	939	11 R04104	puri gene product
45	30	66.7	943	6 P50623	Partial sequence o

ALIGNMENTS

RESULT 1
ID W94488 standard; peptide: 9 AA.
XX W94488:
XX 21-APR-1999 (first entry)
DT
XX Human adult brain zonulin N-terminal peptide.
DE
XX Zonulin, mammalian tight junction; zonula occludens toxin; ZOT;
KW Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody;
KW Intestinal mucosa; nasal mucosa; blood brain barrier.
XX
OS Homo sapiens.
PN W09852415-A1.
XX
PD 26-NOV-1998.
XX
PF 28-APR-1998: 98WO-US07636.
XX
PR 21-MAY-1997: 97US-0859931.
XX
PA (UYMA-) UNTV MARYLAND BALTIMORE.
XX
PI Fasano A:
XX WPI: 1999-070123/06.
XX
XX New purified zonulin - which is capable of reversibly opening
PT mammalian tight junctions, used for enhancing the delivery of agents
PT across intestinal and nasal mucosa and blood brain barrier

PS Claim 2; Page 45; 64pp; English.

XX The present invention describes pure zonulin which has an apparent
CC molecular weight of 47 kD, as determined by SDS-PAGE, which is
CC recognised by both anti-tau polyclonal antibody and by anti-zonula
CC occludens toxin (ZOT) polyclonal antibody, and is capable of reversibly
CC opening mammalian tight junctions. Zonulin proteins function as
CC physiological modulators of mammalian tight junctions. They can be used
CC for enhancing the absorption of therapeutic agents across tight
CC junctions of intestinal and nasal mucosa and across tight junctions of
CC the blood brain barrier. Zonulin can be used with agents such as drugs,
CC e.g. lidocaine, adenosine, dobutamine, dopamine, epinephrine,
CC norepinephrine, phenotolamine, doxapram, alfentanil, dezocin, nalbuphine,
CC buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurine,
CC mivacurium, succinylcholine, cytarabine, mitomycin doxorubicin,
CC vincristine, vinblastine, methicillin, mezlocillin, piperacillin,
CC cefoxitin, cefepime, cefmetazole and aztreonam, a hormone e.g.
CC testosterone, nandrolone, menotropins, insulin, urofollitropin,
CC interferon-alpha, interferon-beta, interferon-gamma, interleukin-1
CC (IL-1), IL-2, IL-4, IL-8, polyvalent IgG, specific IgG, IgA, or IgM.
CC The proteins can also be used for the production of antibodies which can
CC be used to assay for zonulin in body tissue or fluids, or in affinity-
CC purification of zonulin. The present sequence represents an N-terminal
CC peptide of zonulin.

XX
SQ Sequence 9 AA:

Query Match 100.0%; Score 45; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTFYTDAYS 9
Db 1 vtfytdavs 9

RESULT 2
Y79131
ID Y79131 standard; Peptide; 9 AA.
XX
AC Y79131;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human adult brain zonulin N-terminal sequence.
XX
KW Zonulin; antagonist; zonula occludens toxin receptor;
KW human; blood-brain barrier; antiinflammatory;
KW gastrointestinal inflammation; therapy.
XX
OS Homo sapiens.
XX
PN WO200007609-A1.
XX
PD 17-FEB-2000.
XX
PE 28-JUL-1999; 99WO-US16663.
XX
PR 03-AUG-1998; 98US-0127815.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Fasano A;
XX
DR WPI; 2000-205565/18.
XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
XX
PS Example 3; Fig 6; 69pp; English.
XX

CC The present sequence is that of the N-terminal region of adult
CC human brain zonulin. The N-terminal sequences of human adult and
CC fetal zonulins (see Y79130-36) were compared with Vibrio cholerae
CC zonula occludens toxin (ZOT) to identify a common motif thought
CC to be involved in receptor binding. Peptide antagonists (see
CC Y79105-29) based on this motif are useful as antiinflammatory
CC agents for treatment of gastrointestinal inflammation, and for
CC treatment of conditions associated with breakdown of the blood-brain
CC barrier.

XX
SQ Sequence 9 AA:

Query Match 100.0%; Score 45; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTFYTDAYS 9
Db 1 vtfytdavs 9

RESULT 3
Y41173
ID Y41173 standard; protein; 205 AA.
XX
AC Y41173;
XX
DT 31-JAN-2000 (first entry)
XX
DE Llama Vhh polypeptide hybrid sequence.

XX
KW Lymphocyte; multiple cell surface; antigen; ligand; T cell response;
KW B cell response; immune enhancement; immunosuppression; immunotherapy;
KW cancer; infectious disease; multiple surface receptor; AIDS;
KW hypersensitivity; vascular disease; transplant rejection; Vhh protein.

XX
OS Llama llama.
XX
PN WO9942077-A2.
XX
PD 26-AUG-1999.

XX
PE 18-FEB-1999; 99WO-US03309.
XX
PR 19-FEB-1998; 98US-0075274.
PR 16-NOV-1998; 98US-0108683.
XX
PA (XCYT-) XCYTE THERAPIES INC.

XX
PI Ledbetter JA, Hayden Ledbetter M, Brady WA, Grosmaire LS, Law C;
PI Dua R;
XX

DR WPI; 1999-633594/54.

XX
PT Regulation of lymphocyte activation using, e.g. multispecific
PT molecules, used for treatment of cancer -
XX
PS Claim 43; Fig 16A; 114pp; English.

XX The invention relates to a method of regulating lymphocyte activation
CC by selectively binding multiple cell surface (mcs) antigens expressed by
CC the same lymphocyte. The method and the products are used to regulate
CC lymphocyte activation. Lymphocytes can be incubated with immobilized
CC ligands or Ab's or their fragments specific for the target Ag's in order
CC to achieve Ag aggregation in vitro. The multispecific molecules that
CC contain multiple binding specificities in a single soluble molecule are
CC especially useful in aggregating multiple Ag's in vivo resulting in
CC lymphocyte activation. Multispecific molecules may also be constructed
CC to inhibit lymphocyte activation by blocking delivery of activation
CC signals to the cells. The methods and the products are useful for
CC regulation of T and B cell responses in vivo and in vitro. The activation
CC of signals may result in either immune enhancement or immunosuppression.

CC Expanded T and/or B cells are used in adoptive immunotherapy of cancer
 CC and infectious diseases such as AIDS. Aggregation of multiple surface
 CC receptors or inhibition of lymphocyte activation is not only limited to
 CC treatment of, e.g. immunodeficiency, infectious diseases and cancer and
 CC can also be applied for a wider variety of treatments such as for
 CC suppression of autoimmunity, hypersensitivity, vascular disease and
 CC transplant rejection. Also, stimulation of human T cells with immobilized
 CC Ab's specific for three T cell surface Ag's results in enhanced
 CC proliferation when compared with stimulation by two immobilized Ab's.
 CC Sequences Y41165-73 represent Llana Vhh polypeptide unique hybrid
 CC sequences Y41174-79 represent Llana Vhh complete sequences and Y41180-85
 CC represent Llana constant region sequences.

CC Sequence 205 AA;

Query Match 75.6%; Score 34; DB 20; Length 205;

Best Local Similarity 85.7%; Pred. No. 39;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 TRPTDAV 8

DB 58 tftydsy 64

RESULT 4

R08221 R08221 standard; protein; 1475 AA.

AC R08221;

DT 06-MAR-1991 (first entry)

DE Recombinant alpha amylase pullulanase enzyme.

KM Starch; pullulan; saccharification; debranching; liquefaction;

OS hydrolysis.

XX Clostridium thermohydrosulphuricum DSM 3783.

XX Key Location/Qualifiers

XX Peptide 1..31

XX EP402092-A.

XX PD 12-DEC-1990.

XX PF 05-JUN-1990; 90BP-0306088.

XX PR 05-JUN-1989; 89US-0361368.

XX PA (ALKO-) ALKO LTD.

XX PI Melasniemi H, Paloheimo M;

XX DR WPI, 1990-370103/50.

XX PT P-PSDB; R08221.

XX PT Heat stable enzyme with both alpha-amylase and pullulanase

XX PT activities - prepd. by expressing Clostridium thermohydrosulphuricum

XX PT DNA in host cells, useful in hydrolysis of starch, etc.

XX PS Claim 6; Fig 5; 127pp; English.

CC The recombinant enzyme was produced by cloning the gene encoding it
 CC from a C. thermohydrosulphuricum genomic library and using it to
 CC express the enzyme in a host cell e.g. E.coli which is easier to
 CC cultivate and has less complex nutritional requirements than C.
 CC thermohydrosulphuricum (which is an obligate anaerobic thermophilic
 CC organism). The enzyme is used to hydrolyse starch, amylase or
 CC pullulan, opt. in conjunction with a glucogenic or maltogenic enzyme.
 CC It has an optimum temp. range of 80-85 deg.C, i.e. 5 deg. lower than

CC that of the native enzyme, but the heat stability is the same. At
 CC least 10 differently sized polypeptides having the same enzymic
 CC activities were produced, of Mr 100 000 - 165 000.

CC Sequence 1475 AA;

Query Match 75.6%; Score 34; DB 11; Length 1475;

Best Local Similarity 66.7%; Pred. No. 3.2e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VTFPTDAVS 9

DB 222 itfydsvs 230

RESULT 5

Y41166 Y41166 standard; protein; 183 AA.

AC Y41166;

DT 31-JAN-2000 (first entry)

DE Llana Vhh polypeptide hybrid sequence.

KM Lymphocyte; multiple cell surface; antigen; ligand; T cell response;

KM B cell response; immune enhancement; immunosuppression; immunotherapy;

KM cancer; infectious disease; multiple surface receptor; AIDS;

KM hypersensitivity; vascular disease; transplant rejection; Vhh protein.

XX Llana Llana.

XX PN WO942077-A2.

XX PD 26-AUG-1999.

XX PF 18-FEB-1999; 99WO-US03309.

XX PR 19-FEB-1998; 98US-0075274.

XX PR 16-NOV-1998; 98US-0108683.

XX PA (XCYT-) XCYTE THERAPIES INC.

XX PI Ledbetter JA, Hayden Ledbetter M, Brady WA, Grosmaire LS, Law C;

XX PI Dua R;

XX DR WPI, 1999-633594/54.

XX PT Regulation of lymphocyte activation using, e.g. multispecific

XX PT molecules, used for treatment of cancer -

XX PS Claim 43; Fig 16A; 114pp; English.

CC The invention relates to a method of regulating lymphocyte activation
 CC by selectively binding multiple cell surface (mcs) antigens expressed by
 CC the same lymphocyte. The method and the products are used to regulate
 CC lymphocyte activation. Lymphocytes can be incubated with immobilized
 CC ligands or Ab's or their fragments specific for the target Ag's in order
 CC to achieve Ag aggregation in vitro. The multispecific molecules that
 CC contain multiple binding specificities in a single soluble molecule are
 CC especially useful in aggregating multiple Ag's in vivo resulting in
 CC lymphocyte activation. Multispecific molecules may also be constructed
 CC to inhibit lymphocyte activation by blocking delivery of activation
 CC signals to the cells. The methods and the products are useful for
 CC regulation of T and B cell responses in vivo and in vitro. The activation
 CC of signals may result in either immune enhancement or immunosuppression.
 CC Expanded T and/or B cells are used in adoptive immunotherapy of cancer
 CC and infectious diseases such as AIDS. Aggregation of multiple surface
 CC receptors or inhibition of lymphocyte activation is not only limited to
 CC treatment of, e.g. immunodeficiency, infectious diseases and cancer and
 CC can also be applied for a wider variety of treatments such as for
 CC suppression of autoimmunity, hypersensitivity, vascular disease and

CC transplant rejection. Also, stimulation of human T cells with immobilized
 CC Ab's specific for three T cell surface Ag's results in enhanced
 CC proliferation when compared with stimulation by two immobilized Ab's.
 CC Sequences Y41165-73 represent llama Vhh polypeptide unique hybrid
 CC sequences. Y41174-79 represent llama Vhh complete sequences and Y41180-85
 CC represent llama constant region sequences.

XX Sequence 183 AA;

Query Match 71.1%; Score 32; DB 20; Length 183;

Best Local Similarity 62.5%; Pred. No. 84; Mismatches 1; Indels 0; Gaps 0;

YY 1 VTFYTDAY 8
 :||| :| :|
 Db 55 itfyadv 62

RESULT 6

R47450 ID R47450 standard; Protein; 477 AA.

XX AC R47450;

XX DT 24-JUN-1994 (first entry)

XX DE T84.12 Heavy chain.

XX KM Chimeric; carcinoembryonic antigen; CCA; murine; mouse; constant;

XX KW region; transform; myeloma cell; light chain; tumour.

XX OS Synthetic.

XX PN WO9325237-A.

XX PD 23-DEC-1993.

XX PF 15-JUN-1993; 93WO-US05709.

XX PR 15-JUN-1992; 92US-0904074.

XX PA (CITY) CITY OF HOPE.

XX PI Fischer R, Paxton R, Shively JE, Wu A, Yang YHJ;

XX PI Yang YH;

XX DR WPI: 1994-007204/01.

XX DR N-PSDB; Q54652.

XX PT New chimeric T84.12 antibody active against carcinoembryonic

XX PT antigen - has murine variable and human constant regions, also

XX PT DNA encoding it and transformed myeloma cells

XX PS Claim 1; Page 17; 27pp; English.

XX CC The sequences (Q54651-52) show the light and heavy chain cDNAs

XX CC of murine T84.12. The T84.12 antibody is directed against the

XX CC tumour marker carcinoma embryonic antigen, and is useful for

XX CC tumour imaging and immunotherapy.

XX CC The amino acid sequence given in the specification has been

XX CC incorrectly identified as a nucleic acid sequence, therefore

XX CC unacceptable characters have been represented as an 'N'.

XX CC The amino acid sequence given below has been derived from the

XX CC cDNA, by the indexer.

XX Sequence 477 AA;

Query Match 71.1%; Score 32; DB 15; Length 477;

Best Local Similarity 62.5%; Pred. No. 2.3e+02; Mismatches 1; Indels 0; Gaps 0;

YY 1 VTFYTDAY 8
 :||| :| :|
 Db 84 itfyadv 91

RESULT 7

R47453 ID R47453 standard; Protein; 477 AA.

XX AC R47453;

XX DT 24-JUN-1994 (first entry)

XX DE chIT84.12 H3 heavy chain.

XX KM Chimeric; carcinoembryonic antigen; CCA; murine; mouse; constant;

XX KW region; transform; myeloma cell; light chain; tumour.

XX OS Synthetic.

XX PN WO9325237-A.

XX PD 23-DEC-1993.

XX PF 15-JUN-1993; 93WO-US05709.

XX PR 15-JUN-1992; 92US-0904074.

XX PA (CITY) CITY OF HOPE.

XX PI Fischer R, Paxton R, Shively JE, Wu A, Yang YHJ;

XX PI Yang YH;

XX DR WPI: 1994-007204/01.

XX DR N-PSDB; Q54655.

XX PT New chimeric T84.12 antibody active against carcinoembryonic

XX PT antigen - has murine variable and human constant regions, also

XX PT DNA encoding it and transformed myeloma cells

XX PS Claim 1; Page 22-23; 27pp; English.

XX CC The sequences (Q54651-52) show the light and heavy chain cDNAs

XX CC of murine T84.12. The T84.12 antibody is directed against the

XX CC tumour marker carcinoma embryonic antigen, and is useful for

XX CC tumour imaging and immunotherapy.

XX CC The amino acid sequence given in the specification has been

XX CC incorrectly identified as a nucleic acid sequence, therefore

XX CC unacceptable characters have been represented as an 'N'.

XX CC The amino acid sequence given below has been derived from the

XX CC cDNA, by the indexer.

XX Sequence 477 AA;

Query Match 71.1%; Score 32; DB 15; Length 477;

Best Local Similarity 62.5%; Pred. No. 2.3e+02; Mismatches 1; Indels 0; Gaps 0;

YY 1 VTFYTDAY 8
 :||| :| :|
 Db 84 itfyadv 91

RESULT 8

W13514 ID W13514 standard; Protein; 119 AA.

XX AC W13514;

XX DT 28-OCT-1997 (first entry)

```

XX DE Anti-melanoma antibody heavy chain clone G57.
XX KM Human; monoclonal antitumour antibody; peripheral blood lymphocyte;
XX KW cancer; tumorigenesis; anticancer vaccine.
XX OS
XX FH Homo sapiens.
XX FT Key Location/Qualifiers
XX FT Region 31..35
XX FT /Label= CDRI
XX FT Region 50..66
XX FT /Label= CDR2
XX FT Region 98..108
XX FT /Label= CDR3
XX FT
XX PD WO9702479-A2.
XX PD 23-JAN-1997.
XX PF 28-JUN-1996; 96WO-IB01032.
XX PR 30-JUN-1995; 95US-0497647.
XX PA (UYVA ) UNIV YALE.
XX PI Cal X, Garen A;
XX PI
XX DR WPI; 1997-109061/10.
XX PT Prodn. of human monoclonal anti-tumour antibodies - by screening a
XX PT fusion phase library produced using peripheral blood lymphocytes
XX PT from a cancer patient
XX PS Claim 19; Page 52; 82pp; English.
XX CC A process for isolating and synthesising human monoclonal anti-tumour
XX CC antibodies has been produced. The process involves: (a) constructing at
XX CC least one fusion phase library from the peripheral blood lymphocytes
XX CC (PBs) of a cancer patient; (b) screening for anti-tumour antibodies in
XX CC the phage library in a binding assay with cultured tumour cells of the
XX CC same type as the patient's tumour; (c) removing extraneous antibodies by
XX CC absorption against normal human cells; (d) cloning the phage selected in
XX CC step (b) and (c); (e) assaying the specificity of the cloned phage by
XX CC incubating the phage with at least two types of cultured normal cells;
XX CC and (f) further testing the specificity of cloned phage that do not bind
XX CC to either cell line of cultured normal cells in further binding assays
XX CC to cultured tumour cells derived from more than one other tumour that is
XX CC not the patient's tumour. The present sequence represents a human heavy
XX CC chain antibody, from an scFv antibody fusion phase library, produced by
XX CC a method as described above. The antibodies produced can be used for
XX CC diagnostic and therapeutic applications and for isolating tumour
XX CC antigens for studying tumorigenesis or for use as anti-cancer vaccines.
XX CC The human antibodies have low immunogenicity in humans compared to
XX CC murine monoclonal antibodies (MAbs). Since the antibodies are isolated
XX CC from fusion phase libraries, their affinity and specificity for a
XX CC tumour cell line can be improved by genetic manipulations.
XX SQ Sequence 119 AA:

Query Match 68.9%; Score 31; DB 18; Length 119;
Best Local Similarity 71.4%; Pred. No. 82;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 2 TFFYTDAY 8
Db 58 tytydsv 64

RESULT 9
Y41168
ID Y41168 standard; protein: 208 AA.

```

```

XX AC Y41168;
XX DT 31-JAN-2000 (first entry)
XX DE Llama Vhh polypeptide hybrid sequence.
XX KW Lymphocyte; multiple cell surface; antigen; ligand; T cell response;
XX KW B cell response; immune enhancement; immunosuppression; immunotherapy;
XX KW cancer; infectious disease; multiple surface receptor; AIDS;
XX KW hypersensitivity; vascular disease; transplant rejection; Vhh protein.
XX OS Llama llama.
XX PN WO942077-A2.
XX PD 26-AUG-1999.
XX PF 18-FEB-1999; 99WO-0503309.
XX PR 19-FEB-1998; 98US-0075274.
XX PR 16-NOV-1998; 98US-0108683.
XX PA (XCYT-) XCYTE THERAPIES INC.
XX PI Ledbetter JA, Hayden Ledbetter M, Brady WA, Grosmaire LS, Law C;
XX PI Dua R;
XX DR WPI; 1999-633594/54.
XX PT Regulation of lymphocyte activation using, e.g. multispecific
XX PT molecules, used for treatment of cancer -
XX PS Claim 43; Fig 16A; 114pp; English.
XX CC The invention relates to a method of regulating lymphocyte activation
XX CC by selectively binding multiple cell surface (mcs) antigens expressed by
XX CC the same lymphocyte. The method and the products are used to regulate
XX CC lymphocyte activation. Lymphocytes can be incubated with immobilized
XX CC ligands or Ab's or their fragments specific for the target Ag's in order
XX CC to achieve Ag aggregation in vitro. The multispecific molecules that
XX CC contain multiple binding specificities in a single soluble molecule are
XX CC especially useful in aggregating multiple Ag's in vivo resulting in
XX CC lymphocyte activation. Multispecific molecules may also be constructed
XX CC to inhibit lymphocyte activation by blocking delivery of activation
XX CC signals to the cells. The methods and the products are useful for
XX CC regulation of T and B cell responses in vivo and in vitro. The activation
XX CC of signals may result in either immune enhancement or immunosuppression.
XX CC Expanded T and/or B cells are used in adoptive immunotherapy of cancer
XX CC and infectious diseases such as AIDS. Aggregation of multiple surface
XX CC receptors or inhibition of lymphocyte activation is not only limited to
XX CC treatment of, e.g. immunodeficiency, infectious diseases and cancer and
XX CC can also be applied for a wider variety of treatments such as for
XX CC suppression of autoimmunity, hypersensitivity, vascular disease and
XX CC transplant rejection. Also, stimulation of human T cells with immobilized
XX CC Ab's specific for three T cell surface Ag's results in enhanced Ab's.
XX CC Proliferation when compared with stimulation by two immobilized Ab's.
XX CC Sequences Y41165-73 represent Llama Vhh polypeptide unique hybrid
XX CC sequences, Y41174-79 represent Llama Vhh complete sequences and Y41180-85
XX CC represent Llama constant region sequences.
XX SQ Sequence 208 AA:

Query Match 68.9%; Score 31; DB 20; Length 208;
Best Local Similarity 71.4%; Pred. No. 15e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 2 TFFYTDAY 8
Db 58 tytydsv 64

```

```

RESULT 10
ID Y49069 standard; Protein; 457 AA.
XX
AC Y49069;
XX
DT 05-JAN-2000 (first entry)
XX
DE Amino acid sequence encoded by partial dnaB gene.
XX
KM Gram positive bacteria; dnaB; dnaX; dnaB; polC; dnaN; dnaG; helicase;
KW alpha subunit; DNA polymerase III holoenzyme; gamma subunit; tau subunit;
KM clamp loader; glue protein; replication; antibiotic.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 441 /note="Encoded by ACAGGC"
XX
PN WO9337661-A1.
XX
PD 29-JUL-1999.
XX
PF 25-JAN-1999; 99WO-US01547.
XX
PR 27-JAN-1998; 98US-0074522.
XX
PR 22-JUL-1998; 98US-0093727.
XX
PA (UVRQ ) UNIV ROCKEFELLER.
XX
PI O'Donnell ME, Zhang D, Whipple R;
XX
DR WPI; 1999-590685/50.
XX
DR N-PSDB; 231003.
XX
PT New isolated dnaE, dnaX and dnaB genes from Gram positive bacteria,
XX
PT used to develop screening assays for identifying antibiotic compounds
XX
PI
XX
PS Claim 39; Page 21-25; 132pp; English.
XX
CC This is the amino acid sequence encoded by the partial dnaB gene (231003)
CC of Staphylococcus aureus. The invention relates to a number of isolated
CC DNA molecules from Gram positive bacterium, corresponding to dnaE
CC (231001), dnaX (231002), and dnaB (231003). The polC, dnaN and dnaG genes
CC (231004-231006) are also identified. The dnaE gene corresponds to the
CC alpha subunit of the Escherichia coli, DNA polymerase III holoenzyme,
CC dnaX corresponds to the gamma and tau subunits, and dnaB corresponds to
CC the helicase. The alpha subunit is the actual DNA polymerase, the gamma
CC complex forms the clamp loader and tau is a "glue protein". dnaE encodes
CC both gamma and tau, tau is the product of the full gene, while gamma is
CC the product of the first two thirds of the gene. The DNA sequences of the
CC invention can be used to identify agents that inhibit or promote DNA
CC replication by acting on various parts of the gram positive bacterial DNA
CC polymerase holoenzyme. The products and methods of the invention can be
CC used for identifying pharmacological agents or lead compounds for agents
CC active at the level of a replication protein function, particularly DNA
CC replication. The agents identified can be used as antibiotics.
XX
SQ Sequence 457 AA;

```

```

Query Match 68.9%; Score 31; DB 20; Length 457;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 VTFTDAVS 9
   1 :||| ||
Db 102 vgytldivs 110

```

```

RESULT 11

```

```

Y31769
ID Y31769 standard; Protein; 466 AA.
XX
AC Y31769;
XX
DT 06-DEC-1999 (first entry)
XX
DE Staphylococcus aureus replicative helicase dnaB.
XX
KM Replicative helicase; dnaB; antibacterial; antibiotic; screening;
KW infection; bacteraemia; sepsis; therapy; diagnosis; vaccine.
XX
OS Staphylococcus aureus.
XX
PN WO9946275-A1.
XX
PD 16-SEP-1999.
XX
PF 10-MAR-1999; 99WO-US05286.
XX
PR 12-MAR-1998; 98US-0038909.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI May EM, Earnshaw DL, McDevitt D;
XX
DR WPI; 1999-561656/47.
XX
DR N-PSDB; X87985.
XX
PT Novel replicative helicase dnaB polynucleotides and polypeptides of
XX
PT Staphylococcus aureus used to screen for antibacterial compounds
XX
PI
XX
PS Claim 1; Page 46-47; 53pp; English.
XX
CC This sequence represents the replicative helicase, dnaB, of
CC Staphylococcus aureus strain WCUH 29 (NCIMB 40771). The sequence
CC was deduced from an isolated polynucleotide (see X87985) and shows
CC structural homology to other proteins of the replicative helicase
CC family. The invention provides dnaB polynucleotides and
CC polypeptides, and methods for producing such polypeptides by
CC recombinant techniques. Also provided are methods for utilizing
CC dnaB polypeptides to screen for antibacterial compounds, especially
CC those that target antibiotic-resistant strains. The polypeptides
CC and polynucleotides may be employed as research reagents for the
CC discovery of treatments and diagnostics, particularly for human
CC diseases. The polypeptides can be used to produce antibodies.
CC In vaccine formulations, and to identify agonists and antagonists.
CC These are used to prevent, inhibit or treat diseases, particularly
CC of Helicobacter pylori infections, such as gastrointestinal
CC carcinoma, and also gastric ulcers and gastritis. The polypeptides
CC can also be used to treat wounds and in-dwelling devices to prevent
CC bacterial adhesion and infection, and to block dnaB protein-mediated
CC mammalian cell invasion.
XX
SQ Sequence 466 AA;

```

```

Query Match 68.9%; Score 31; DB 20; Length 466;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 VTFTDAVS 9
   1 :||| ||
Db 102 vgytldivs 110

```

```

RESULT 12
Y70133
ID Y70133 standard; Protein; 466 AA.
XX
AC Y70133;
XX

```

```

DT 06-JUN-2000 (first entry)
XX Staphylococcus aureus replicative DNA helicase DnaB.
DE
XX
XX Vaccine: antibacterial; prevention; attenuation; detection;
KM Staphylococcal infection; neonatal conjunctivitis; skin infection;
KM toxic shock syndrome; osteomyelitis; DnaB; replicative DNA helicase.
XX
XX Staphylococcus aureus.
OS
XX
XX Key Location/Qualifiers
XX Domain 155..158
XX Domain /label= Antigenic_epitope
XX Domain 178..181
XX Domain /label= Antigenic_epitope
XX Domain 304..306
XX Domain /label= Antigenic_epitope
XX Domain 401..403
XX Domain /label= Antigenic_epitope
XX Domain 405..408
XX Domain /label= Antigenic_epitope
XX Domain 411..416
XX Domain /label= Antigenic_epitope
XX
XX MO200012678-A2.
XX
XX 09-MAR-2000.
XX
XX 31-AUG-1999; 99WO-US19726.
XX
XX 01-SEP-1998; 98US-0098964.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Bailey CC, Choi GH;
XX
XX WPI: 2000-237864/20.
XX
XX N-PSDB: 251218.
XX
XX Staphylococcus aureus polypeptide useful for preventing or attenuating
PT a Staphylococcal infection comprises one of 32 sequences of 100-1277
XX amino acids or their fragments.
XX
XX Claim 9; Page 12; 14pp; English.
XX
XX The present sequence is a replicative DNA helicase DnaB
CC from Staphylococcus aureus genomic DNA library.
CC The present sequence is useful in preparation of vaccines for
CC prevention or attenuation of Staphylococcal infections (especially
CC S. aureus infections) which may cause conditions such as neonatal
CC conjunctivitis, osteomyelitis, skin infections and toxic shock syndrome.
CC The present sequence is also useful for detecting Staphylococcal
CC infections in biological samples.
XX
XX Sequence 466 AA;
SQ

```

Query Match 68.9%; Score 31; DB 21; Length 466;
Best Local Similarity 66.7%; Pred. NO. 3.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 VTFTYDAVS 9
DB 102 vqyytdlvs 110

```

RESULT 13
B33998
ID B33998 standard; Peptide: 37 AA.
XX
AC B33998;
XX
DT 02-FEB-2001 (first entry)
XX

```

XX
DE Human secreted protein encoded by cDNA #39.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KM vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KM neurological disease; infection; human; secreted protein.
XX
XX Homo sapiens.
OS
XX
XX MO200056765-A1.
XX
XX 28-SEP-2000.
XX
XX 16-MAR-2000; 2000WO-US06823.
XX
XX 19-MAR-1999; 99US-0125364.
XX
XX 08-DEC-1999; 99US-0169623.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI: 2000-602215/57.
XX
XX N-PSDB: C59430.
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers.
XX
XX Claim 11; Page 370; 410pp; English.
XX
XX Sequences B33963-B34006 represent the amino acid sequences of 48
CC human secreted proteins encoded by the genes C59392-C59439. The genes
CC and proteins are useful for preventing, ameliorating or treating
CC medical conditions, e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and
CC ulcerative colitis; (c) cardiovascular disorders such as myocardial
CC ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral
CC anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections.
XX
XX Sequence 37 AA;
SQ

```

Query Match 66.7%; Score 30; DB 21; Length 37;
Best Local Similarity 66.7%; Pred. NO. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 VTFTYDAVS 9
DB 18 llytdlvs 26

```

RESULT 14
Y41320
ID Y41320 standard; Protein: 48 AA.
XX
AC Y41320;
XX
DT 02-DEC-1999 (first entry)
XX
DE Human secreted protein encoded by gene 13 clone HMSK186.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX

KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asplenia; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9947540-A1.
 XX
 PD 23-SEP-1999.
 XX
 PF 18-MAR-1999; 99MO-US05804.
 XX
 PR 19-MAR-1998; 98US-0078563.
 PR 19-MAR-1998; 98US-0078566.
 PR 19-MAR-1998; 98US-0078573.
 PR 19-MAR-1998; 98US-0078574.
 PR 19-MAR-1998; 98US-0078576.
 PR 19-MAR-1998; 98US-0078577.
 PR 19-MAR-1998; 98US-0078578.
 PR 19-MAR-1998; 98US-0078579.
 PR 19-MAR-1998; 98US-0078581.
 PR 01-APR-1998; 98US-0080312.
 PR 01-APR-1998; 98US-0080313.
 PR 01-APR-1998; 98US-0080314.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
 PI Mei Y, Endress GA, Duan KD, Kyaw H, Ebner R, Lafleur DW;
 PI Olsen HS, Shi Y, Moore PA;
 DR WPI: 1999-562050/47.
 DR N-PSDB: Z24823.
 XX
 PT New isolated human genes, useful for diagnosis and treatment of e.g.
 PT cancers, neurological disorders, immune diseases, inflammation or blood
 PT disorders
 XX
 PS Claim 11: Page 365; 484pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. Z24802) for increasing the stability of the fused protein
 CC as compared to the human protein only.
 CC The invention relates to 95 novel genes and their fragments (nucleic
 CC acid sequences: Z24811-Z24907; amino acid sequences Y41308-Y41404) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 95 polynucleotides, based on
 CC which tissues they are most highly expressed in (see Z24811 for described
 CC uses).
 CC
 SQ Sequence 48 AA:

Query Match 66.7%; Score 30; DB 20; Length 48;
 Best Local Similarity 71.4%; Pred. No. 48;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VFPTDA 7
 |||:
 Db 26 vlfids 32

RESULT 15
 Y02500

ID Y02500 standard; Protein; 56 AA.
 XX
 AC Y02500;
 XX
 DR 15-JUL-1999 (first entry)
 XX
 DE Clone selected after panning a NK library of the invention.
 XX
 KW Screening; functional polypeptide; ligand; non-functional;
 KW enrichment; single chain antibody; SCFv.
 XX
 OS Synthetic.
 XX
 PN WO9920749-A1.
 XX
 PD 29-APR-1999.
 XX
 PF 20-OCT-1998; 98MO-GB03135.
 XX
 PR 21-NOV-1997; 97US-0066729.
 PR 20-OCT-1997; 97GB-0022131.
 PR 13-NOV-1997; 97US-0065428.
 XX
 PA (MED-) MEDICAL RES COUNCIL.
 XX
 PI Tomlinson I, Winter G;
 XX
 DR WPI: 1999-288302/24.
 XX
 PT Screening for functional polypeptides which bind a ligand
 XX
 PS Example 3; Fig 4; 67pp; English.
 XX
 CC The specification describes a method for screening for functional
 CC polypeptides which bind a ligand. The method comprises contacting a
 CC repertoire of polypeptides with a generic ligand, and then screening
 CC selected functional polypeptides with a target ligand. The method
 CC permits the removal from a chosen repertoire of polypeptides, those
 CC which are non-functional, e.g. as a result of the introduction of
 CC frame-shift mutations, stop codons, folding mutants or expression
 CC mutants which would be or are incapable of binding to any target
 CC ligand. The method also permits the enrichment of a chosen repertoire
 CC of polypeptides for those polypeptides which are functional, well folded
 CC and highly expressed. The polypeptides obtained can be used in
 CC diagnostic, prophylactic and therapeutic procedures. Y02473-Y02525
 CC represent clones selected after panning primary and somatic NK
 CC libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
 CC NIP-BSA and hen egg lysozyme).
 CC
 SQ Sequence 56 AA:

Query Match 66.7%; Score 30; DB 20; Length 56;
 Best Local Similarity 62.5%; Pred. No. 57;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VFPTDAV 8
 |||:
 Db 13 vlyadsv 20

Search completed: June 13, 2001, 14:25:46
 Job time: 665 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:11 ; Search time 87.97 Seconds
(without alignments)
7.031 Million cell updates/sec

Title: PCT-US01-05825A-26

Perfect score: 45

Sequence: 1 VTFYTDALS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	36	80.0	260	2	A84066	molbdenum transpo
2	34	75.6	365	2	C83885	hypothetical prote
3	34	75.6	1475	2	A44765	alpha-amylase (EC
4	34	75.6	1481	2	S28669	amylopolyluanase P
5	33	73.3	474	2	B75126	hypothetical prote
6	33	73.3	651	2	T25953	hypothetical prote
7	33	73.3	695	2	T46113	hypothetical prote
8	32	71.1	486	2	C64765	yeast protein precu
9	32	71.1	520	1	A41771	3-oxoacid CoA-tran
10	31	68.9	220	1	D82139	conserved hypotet
11	31	68.9	256	2	H69201	hypothetical prote
12	31	68.9	301	2	S20081	surface virulence
13	31	68.9	338	2	T47218	glyceraldenhyde-3-p
14	31	68.9	553	2	F81282	probable ferredoxi
15	31	68.9	645	2	S49570	penicillin-binding
16	31	68.9	649	2	S74823	N-acetylmutamoyl-L
17	31	68.9	756	2	F83704	homocysteine methyl
18	31	68.9	879	2	E67992	conserved hypotet
19	31	68.9	1376	2	G00043	osteonodogen - hum
20	31	68.9	2247	2	T16637	hypothetical prote
21	30	66.7	73	2	T42302	hypothetical prote
22	30	66.7	124	2	S20784	Ig heavy chain V r
23	30	66.7	216	2	E75425	hypothetical prote
24	30	66.7	240	2	T45727	hypothetical prote
25	30	66.7	254	2	E83619	probable transport
26	30	66.7	260	2	T30236	methyltransferase
27	30	66.7	286	2	B71663	deoxyribonuclease
28	30	66.7	356	3	JC7131	isocitrate dehydro
29	30	66.7	369	2	A39309	

30	30	66.7	385	2	H70937	hypothetical prote
31	30	66.7	387	2	A71960	probable nitrogena
32	30	66.7	387	2	D64547	iron-sulfur cofact
33	30	66.7	438	2	T25235	hypothetical prote
34	30	66.7	452	2	JC6561	UDP-N-acetylmuramo
35	30	66.7	459	2	T11489	NADH dehydrogenase
36	30	66.7	578	2	T11659	hypothetical prote
37	30	66.7	579	2	A70954	hypothetical prote
38	30	66.7	581	2	T23922	hypothetical prote
39	30	66.7	618	2	S11445	parasporeal crystal
40	30	66.7	641	2	T03095	homeoprotein Sail
41	30	66.7	658	2	G81727	conserved hypotet
42	30	66.7	749	2	C72725	hypothetical prote
43	30	66.7	818	2	T01105	disease resistance
44	30	66.7	1155	2	A26513	parasporeal crystal
45	30	66.7	1155	2	JD0002	parasporeal crystal

ALIGNMENTS

RESULT 1
A84066
molbdenum transport system (molybdate-binding protein) BH3329 [imported] - Bacillus
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
C:Accession: A84066
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650, M0ID:20263314
A:Accession: A84066
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BA07048.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3329
C:Superfamily: molybdate-binding periplasmic protein

Query Match 80.0%; Score 36; DB 2; Length 260;

Best Local Similarity 66.7%; Pred. No. 6.8; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 2;

QY 1 VTFYTDALS 9
DB 192 IVFTDALS 200

RESULT 2
C83885
hypothetical protein BH1883 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000

C:Accession: C83885
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650, M0ID:20263314
A:Accession: C83885

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-365 <STO>
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05602.1; GSPDB:G

A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1883

Query Match 75.6%; Score 34; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VFYTDVAVS 9
 |||||
 Db 138 VFYLDGDS 146

RESULT 3

A44765
 alpha-amylase (EC 3.2.1.1) / alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) prec
 C:Species: Thermomicrobacter thermophilus
 C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 15-Oct-1999
 C:Accession: A44765
 R:Melasniemi, H.; Palohelmo, M.; Hemioe, L.
 J. Gen. Microbiol. 136, 447-454, 1990
 A:Title: Nucleotide sequence of the alpha-amylase-pullulanase gene from Clostridium ther
 A:Reference number: A44765; MUID:90362027
 A:Accession: A44765
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1475 <MEI>
 A:Cross-references: EMBL:M28471; NID:g144726; PIDN:AAA23205.1; PID:g144727
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: fibronectin type III repeat homology
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 75.6%; Score 34; DB 2; Length 1475;
 Best Local Similarity 66.7%; Pred. No. 97;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VFYTDVAVS 9
 :|||:|
 Db 222 ITFYDVS 230

RESULT 4

S28669
 amylopullulanase precursor [validated] - Thermomicrobacter ethanolicus (strain 39E)
 N:Contains: alpha amylase (EC 3.2.1.1); alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2
 C:Species: Thermomicrobacter ethanolicus
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 03-Nov-2000
 C:Accession: S28669; A47341
 R:Mathupala, S.; Saha, B.C.; Zeikus, J.G.
 Biochem. Biophys. Res. Commun. 166, 126-132, 1990
 A:Title: Substrate competition and specificity at the active site of amylopullulanase fr
 A:Reference number: S28669; MUID:90147689
 A:Accession: S28669
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1481 <MAT1>
 A:Cross-references: EMBL:M97665; NID:g144719; PIDN:AAA23201.1; PID:g144720
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
 R:Mathupala, S.P.; Lowe, S.E.; Podkovyrov, S.M.; Zeikus, J.G.
 J. Biol. Chem. 268, 16332-16344, 1993
 A:Title: Sequencing of the amylopullulanase (apu) gene of Thermomicrobacter ethanolicus
 A:Reference number: A47341; MUID:93346376
 A:Accession: A47341
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1481 <MAT2>
 A:Cross-references: GB:M97665; NID:g144719; PIDN:AAA23201.1; PID:g144720
 C:Genetics:
 A:Gene: apu
 A:Start codon: GTG
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: alpha-amylase core homology
 C:Keywords: glycosidase; hydrolase; multifunctional enzyme; polysaccharide degradation

F:1-33/Domain: signal sequence #status predicted <SIG>
 F:32-1481/Product: amylopullulanase #status predicted <MAT>
 F:596-737/Domain: alpha-amylase core homology <MAT>

Query Match 75.6%; Score 34; DB 2; Length 1481;
 Best Local Similarity 66.7%; Pred. No. 97;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VFYTDVAVS 9
 :|||:|
 Db 222 ITFYDVS 230

RESULT 5

B75126
 hypothetical protein PAB1825 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: B75126
 R:Anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
 A:Reference number: A75001
 A:Accession: B75126
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-474 <KAW>
 A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CA849723.1; PID:g545
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB1825

Query Match 73.3%; Score 33; DB 2; Length 474;
 Best Local Similarity 75.0%; Pred. No. 50;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TFYTDVAVS 9
 :|||:|
 Db 281 TFYTDLS 288

RESULT 6

T25953
 hypothetical protein ZC204.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T25953
 R:Wamsley, P.; Kramer, J.
 submitted to the EMBL Data Library, December 1996
 A:Description: The sequence of C. elegans cosmid ZC204.
 A:Reference number: Z20116
 A:Accession: T25953
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-651 <WAM>
 A:Cross-references: EMBL:U80839; PIDN:AA837917.1; GSPDB:GN00020; CESP:ZC204.8
 A:Experimental source: strain Bristol N2; clone ZC204
 C:Genetics:
 A:Gene: CESP:ZC204.8
 A:Map position: 2
 A:Intons: 28/2; 291/3; 348/2; 427/2; 516/1; 566/3; 644/2

Query Match 73.3%; Score 33; DB 2; Length 651;
 Best Local Similarity 85.7%; Pred. No. 68;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TFYTDVAVS 8
 :|||:|
 Db 565 TFYTDVAVS 571

```

RESULT 7
T46113
hypothetical protein T27B3.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
R:Nakatani, G.; Fartmann, B.; Dauner, D.; Stert, W.; Holland, R.; Welschgartner, M.;
submitted to the Protein Sequence Database, January 2000
C:Accession: T46113
A:Reference number: 223022
A:Accession: T46113
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-695 <NVA>
A:Cross-references: EMBL:AL137079
A:Experimental source: cultivar Columbia; BAC clone T27B3
C:Genetics:
A:Map position: 3
A:Introns: 20/3; 81/3; 107/2; 144/1; 228/3; 250/3; 272/3; 294/3; 316/3; 338/3; 374/3; 39
A:Note: T27B3.60

Query Match
Best Local Similarity 73.3%; Score 33; DB 2; Length 695;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VFPTDAYS 9
:::|||||
Db 211 LSFYTDKVS 219

RESULT 8
C64765
Yait protein precursor - Escherichia coli
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999
C:Accession: C64765
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: C64765
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-486 <BLAT>
A:Cross-references: GB:AE000144; GB:U00096; NID:91786568; PIDN:AC73474.1; PID:91786569;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: Yait
F:1-28/Domain: signal sequence #status predicted <STC>
F:29-486/Product: Yait protein #status predicted <MAT>

Query Match
Best Local Similarity 71.1%; Score 32; DB 2; Length 486;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FYTDAY 8
|||||
Db 139 FYTDAY 144

RESULT 9
A11771
3-oxoacid CoA-transferase (EC 2.8.3.5) precursor, mitochondrial [validated] - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 26-May-2000
C:Accession: A11771; S27952
R:Lin, T.W.; Bridger, W.A.;
J. Biol. Chem. 267, 975-978, 1992
A:Title: Sequence of a cDNA clone encoding pig heart mitochondrial CoA transferase.
A:Reference number: A11771; MUID:92112838

```

```

A:Accession: A11771
A:Molecule type: mRNA
A:Residues: 1-520 <LIN>
A:Cross-references: EMBL:M80534; NID:g164422; PIDN:AAA31019.1; PID:g164423
A:Experimental source: heart
A:Note: sequence extracted from NCBI backbone (NCBI:P:75613)
A:Note: parts of this sequence, including the amino end of the mature protein, were d
C:Function:
A:Description: EC 2.8.3.5 [validated; MUID:92112838]
C:Superfamily: 3-oxoacid CoA-transferase; 3-oxoacid CoA-transferase alpha chain ho
C:Keywords: CoA-transferase; homodimer; ketone body metabolism; mitochondrion
F:1-39/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:38-274/Domain: 3-oxoacid CoA-transferase alpha chain homology <BACA>
F:40-520/Product: 3-oxoacid CoA-transferase #status experimental <MAT>
F:502-506/Domain: 3-oxoacid CoA-transferase beta chain homology <BACB>
F:344/Active site: Glu #status predicted

Query Match
Best Local Similarity 71.1%; Score 32; DB 1; Length 520;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FYTDAY 8
|||||
Db 42 FYTDAY 47

RESULT 10
D82139
conserved hypothetical protein VC1940 [imported] - Vibrio cholerae (group O1 strain N
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: D82139
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Yamahayvan, J.; Bass, S.; Qin, H.; Dragoli, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833
A:Accession: D82139
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-220 <HEI>
A:Cross-references: GB:AE004269; GB:AE003852; NID:99656466; PIDN:AAF95088.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1940
A:Map position: 1

Query Match
Best Local Similarity 68.9%; Score 31; DB 2; Length 220;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VFPTDAYS 9
|||||
Db 176 VFPTDSIN 184

RESULT 11
H69201
hypothetical protein WH762 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum (strain Delta H)
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: H69201
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T
; Liu, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514
A:Accession: H69201
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

```

A:Molecule type: DNA
A:Residues: 1-256 <MTH>
A:Cross-references: GB:AE000854; GB:AE000666; NID:92621839; PIDN:AA85265.1; PID:9262185
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH762
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1420

Query Match 68.9%; Score 31; DB 2; Length 256;
Best Local Similarity 62.5%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFYTDAY 8
DB 110 ITFYMDSV 117

RESULT 12
S20081
surface virulence factor regulator - Escherichia coli
C:Species: Escherichia coli
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S20081
R:de Haan, L.A.M.; Wiltshaw, G.A.; van der Zeijst, B.A.M.; Gaastra, W.
FEMS Microbiol. Lett. 83, 341-346, 1991
A:Title: The nucleotide sequence of a regulatory gene present on a plasmid in an enterot
A:Reference number: S20081
A:Accession: S20081
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-301 <DEH>
A:Cross-references: EMBL:X60106; NID:948934; PIDN:CAA4700.1; PID:948935
A:Note: the authors translated the codon AAA for residue 13 as Asn
C:Superfamily: fapR protein
C:Keywords: DNA binding; transcription regulation

Query Match 68.9%; Score 31; DB 2; Length 301;
Best Local Similarity 62.5%; Pred. No. 81;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFYTDAY 8
DB 161 ITFYMDKV 168

RESULT 13
T47218
glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
R:Sahni, M.; Kinsey, J.A.
submitted to the EMBL Data Library, April 1996
A:Description: Isolation of the Neurospora crassa glyceraldehyde-3-phosphate dehydrogen
A:Reference number: Z24398
A:Accession: T47218
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-338 <SAH>
A:Cross-references: EMBL:U56397; PIDN:AA80570.1
C:Genetics:
A:Gene: gpd-1
A:Map position: LG IIR
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 68.9%; Score 31; DB 2; Length 338;
Best Local Similarity 85.7%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFYTDAY 7
DB 73 VKFYTDA 79

RESULT 14
F81282
Probable ferredoxin Cj1377c [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C:Accession: F81282
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chl
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912
A:Accession: F81282
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-553 <PAR>
A:Cross-references: GB:AL139078; GB:AL111168; NID:96968723; PIDN:CA873804.1; PID:9696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1377c

Query Match 68.9%; Score 31; DB 2; Length 553;
Best Local Similarity 66.7%; Pred. No. 1,540;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VTFYTDAYS 9
DB 326 LVFYTDAYS 334

RESULT 15
S49570
penicillin-binding protein (spore cortex) spoVD - Bacillus subtilis
N:Alternate names: stage V sporulation protein D
C:Species: Bacillus subtilis
C>Date: 15-Nov-1996 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C:Accession: S49570; S43863; D53292; A47691; D69715; S23913
R:Errington, J.
submitted to the EMBL Data Library, September 1993
A:Reference number: S49570
A:Accession: S49570
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-645 <ERR>
A:Cross-references: EMBL:Z25865; NID:9397893; PIDN:CAA81085.1; PID:9580936
R:Daniel, R.A.; Drake, S.; Buchanan, C.E.; Scholle, R.; Errington, J.
J. Mol. Biol. 235, 209-220, 1994
A:Title: The Bacillus subtilis spoVD gene encodes a mother-cell-specific penicillin-b
A:Reference number: S43863; MUID:94118264
A:Accession: S43863
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-589 <DAN>
A:Cross-references: EMBL:Z25865
R:Yanouri, A.; Daniel, R.A.; Errington, J.; Buchanan, C.E.
J. Bacteriol. 175, 7604-7616, 1993
A:Title: Cloning and sequencing of the cell division gene pbb, which encodes penicil
A:Reference number: A53292; MUID:94064553
A:Accession: D53292
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-69 <YAN>
A:Cross-references: GB:L09703; NID:9304162; PIDN:AA836838.1; PID:9304166
R:Daniel, R.A.; Errington, J.
J. Gen. Microbiol. 139, 361-370, 1993
A:Title: DNA sequence of the mure-murD region of Bacillus subtilis 168.
A:Reference number: A47691; MUID:93171879
A:Accession: A47691

A:Molecule type: DNA
 A:Residues: 595-645 <DA2>
 A:Cross-references: EMBL:215056; MID:940160; PIDN:CAA78766.1; PID:940161
 A:Experimental source: strain 168
 A>Note: sequence extracted from NCBI backbone (NCBIN:125659, NCBI:125660)
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteri
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emerson, P.T.; Ertlan, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 lech, J.; Harwood, C.R.; Henuit, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
 Koeltter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y.M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Nieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seroi
 kreich, M.; Tamakoshi, A.; Tanaka, T.; Terpiltra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033
 A:Accession: D69715
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Residues: 1-645 <KUN>
 A:Molecule type: DNA
 A:Cross-references: GB:299111; GB:AL009126; MID:92633699; PIDN:CAB13390.1; PID:92633888
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: spovD
 A:Map position: 133 (degrees)
 A:Start codon: TTG
 C:Superfamily: penicillin-binding protein 3
 C:Keywords: membrane protein; sporulation

Query Match 68.9%; Score 31; DB 2; Length 645;
 Best Local Similarity 85.7%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VVFYTD 7
 I I I I I
 Db 191 VKFYTD 197

Search completed: June 13, 2001, 14:23:11
 Job time: 742 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:36 ; Search time 51.57 Seconds
(without alignments)
5.978 Million cell updates/sec

Title: PCT-US01-05825A-26
Perfect score: 45
Sequence: 1 VTFTDANS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	75.6	1475	1 APU_THEY	P16950 t amylopull
2	34	75.6	1481	1 APU_THEY	P38939 t amylopull
3	32	71.1	314	1 APL_SERCA	P54864 serinus can
4	32	71.1	486	1 YAT_ECOLI	P77199 escherichia
5	32	71.1	520	1 SCOT_PIG	Q29551 sus scrofa
6	31	68.9	301	1 CSVR_ECOLI	P43460 escherichia
7	31	68.9	338	1 G3P_NEOCR	P54118 neuropeptid
8	31	68.9	645	1 SP5D_BACSU	Q03524 bacillus su
9	31	68.9	756	1 METE_BACHD	Q09f61 bacillus ha
10	31	68.9	1375	1 NID2_HUMAN	Q14112 homo sapien
11	30	66.7	369	1 IDH2_YEAST	P28241 saccharomyc
12	30	66.7	385	1 YL90_MYCTU	Q10383 mycobacteri
13	30	66.7	459	1 NUD4_RABIT	Q79436 oryctolagus
14	30	66.7	527	1 ESR2_RABIN	Q9xsb5 bos taurus
15	30	66.7	527	1 ESR2_SHEEP	Q9tuis ovis aries
16	30	66.7	578	1 YDVB_SCHPO	Q14226 schizosach
17	30	66.7	911	1 CIAF_BACTU	P96315 bacillus th
18	30	66.7	1155	1 CIAF_BACTU	P06578 bacillus th
19	30	66.7	1178	1 CIAF_BACTU	P05068 bacillus th
20	30	66.7	1181	1 CIAF_BACTU	Q03748 bacillus th
21	30	66.7	1258	1 ASTN_MOUSE	Q61137 mus musculu
22	30	66.7	1310	1 ASTN_MOUSE	Q14525 homo sapien
23	30	66.7	1325	1 YAB6_SCHPO	Q09847 schizosach
24	29	64.4	105	1 APL_BOVIN	Q77627 bos taurus
25	29	64.4	117	1 HV02_CANFA	P01785 canis fami
26	29	64.4	136	1 YRNE_CAEEL	Q09420 caenorhabd
27	29	64.4	140	1 ATPE_VIBAL	P12988 vibrio algi
28	29	64.4	160	1 YRNE_CAEEL	Q09419 caenorhabd
29	29	64.4	227	1 XNRI_HUMIN	P55334 humicola in
30	29	64.4	296	1 TJUN_AVISI	P05411 avian sarco
31	29	64.4	310	1 APL_CHICK	P18870 gallus galli
32	29	64.4	313	1 APL_COTUA	P12981 coturnix co
33	29	64.4	331	1 APL_HUMAN	P05412 homo sapien

ALIGNMENTS

RESULT 1	APU_THEY	STANDARD:	PRT: 1475 AA.
AC	P16950:		
DT	01-AUG-1990 (Rel. 15, Created)		
DT	01-AUG-1990 (Rel. 15, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	AMYLOPULLULANASE PRECURSOR (ALPHA-AMYLASE/PULLULANASE) [INCLUDES: ALPHA-AMYLASE (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE); PULLULANASE (EC 3.2.1.41) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE)]		
DE	(ALPHA-DEXTRIN ENDO-1,6-ALPHA-GLUCOSIDASE)]		
GN	APU.		
OS	Thermoanaerobacter thermohydrosulfuricus (Clostridium		
OS	thermohydrosulfuricum).		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Thermoanaerobacter group; Thermoanaerobacter.		
OX	NCBI_TaxID=1516;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-E101-69;		
RX	MEDLINE=90362027; PubMed=2391488;		
RA	Melaniemi H., Palohelmo M., Hemioe L.;		
RT	"Nucleotide sequence of the alpha-amylase-pullulanase gene from		
RT	Clostridium thermohydrosulfuricum.";		
RL	J. Gen. Microbiol. 136:447-454(1990).		
RN	[2]		
RP	SEQUENCE OF 32-39.		
RC	SPRAIN-E101-69;		
RX	MEDLINE=88268757; PubMed=3260488;		
RA	Melaniemi H.;		
RT	"Purification and some properties of the extracellular alpha-amylase-		
RT	pullulanase produced by Clostridium thermohydrosulfuricum.";		
RL	Biochem. J. 250:813-818(1988).		
RN	[3]		
RP	FIBRONECTIN TYPE III DOMAINS.		
RX	MEDLINE=93028390; PubMed=1409594;		
RA	Bork P., Doolittle R.F.;		
RT	"Proposed acquisition of an animal protein domain by bacteria.";		
RT	Proc. Natl. Acad. Sci. U.S.A. 89:8950-8954(1992)		
RL	"Catalytic activity: ENDOMYDOLYSIS OF 1,4-ALPHA-GLUCOSIDIC		
CC	LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.		
CC	CATALYTIC ACTIVITY: STARCH-DEBRANCHING ENZYME, HYDROLYZES		
CC	(1-6)-ALPHA-GLUCOSIDIC LINKAGES IN PULLULAN AND STARCH TO		
CC	FORM MALTOSE.		
CC	- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO		
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.		
CC	- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.		
CC	- THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		

34	29	64.4	331	1 APL_PIG	P56432 sus scrofa
35	29	64.4	334	1 APL_MOUSE	P05627 mus musculu
36	29	64.4	334	1 APL_RAT	P17325 rattus norv
37	29	64.4	396	1 NIFS_AZOC	P23120 azotobacter
38	29	64.4	402	1 NIFS_AZOV	P05341 azotobacter
39	29	64.4	409	1 MDPI_PIG	P22412 sus scrofa
40	29	64.4	410	1 MDPI_RABIT	P31429 oryctolagus
41	29	64.4	410	1 MDPI_SHEEP	P43477 ovis aries
42	29	64.4	413	1 SPVA_MOUSE	Q35423 mus musculu
43	29	64.4	414	1 SPVA_RAT	P09139 rattus norv
44	29	64.4	425	1 YLKI_CAEEL	P41949 caenorhabd
45	29	64.4	503	1 IMDH_ARATH	P47996 arabidopsis

EMBL: M97665, AAA23201.1; -;
PIR: S28669; S28669; -;
InterPro: IPR000461; -;
InterPro: IPR001777; -;

SET	DNA_BIND	240	259	BASIC MOTIF.
SET	DOMAIN	263	291	LEUCINE-ZIPPER.
SET	SEQUENCE	314 AA;	34518 MW;	87E936AC37AC2731 CRC64;

Query Match 71.1%; Score 32; DB 1; Length 314;
 Best Local Similarity 75.0%; Pred. No. 26;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 FTYTDAVS 9
 DB 8 TTYEDALS 15

RESULT 4
 YAIT_ECOLI STANDARD; PRT; 486 AA.
 ID YAIT_ECOLI
 AC P71199;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL 52.9 KDA PROTEIN IN HEMB-SBMA INTERGENIC REGION
 DE PRECURSOR.
 GN YAIT.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
 RA Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O., Lew H., Lin D.,
 RA Nameth A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: AE000144; AAC73474.1; -
 CC DR EMBL: U73857; AAB18094.1; -
 CC DR Ecogenet; EC13605; yait.
 CC KW Hypothetical protein; Signal.
 CC FT SIGNAL 1 27
 CC FT CHAIN 28 486
 CC FT SEQUENCE 486 AA; 52856 MW; EE363439B0235EC CRC64;
 CC HYPOTHETICAL PROTEIN YAIT.

Query Match 71.1%; Score 32; DB 1; Length 486;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FTYTDAV 8
 DB 139 FTYTDAV 144

RESULT 5
 SCOT_PIG STANDARD; PRT; 520 AA.
 ID SCOT_PIG
 AC Q29551;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE, MITOCHONDRIAL
 DE PRECURSOR (EC 2.8.3.5) (SUCCINYL COA:3-OXOACID COA-TRANSFERASE).
 GN OXCT OR SCOT.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 40-70 AND 296-305.
 RC TISSUE=Heart;
 RX MEDLINE=92112838; PubMed=1730685;
 RA Lin T., Bridger W.A.;
 RT "Sequence of a cDNA clone encoding pig heart mitochondrial CoA
 RT transferase.";
 RL J. Biol. Chem. 267:975-978(1992).
 RN [2]
 RP ACTIVE SITE.
 RX MEDLINE=94348427; PubMed=7915164;
 RA Rochet J.C., Bridger W.A.;
 RT Identification of glutamate 344 as the catalytic residue in the
 RT active site of pig heart CoA transferase.";
 RL Protein Sci. 3:975-981(1994).
 CC -1- FUNCTION: KEY ENZYME FOR KETONE BODY CATABOLISM. TRANSFERS THE COA
 CC MOIETY FROM SUCCINATE TO ACETOACETATE. FORMATION OF THE ENZYME-COA
 CC INTERMEDIATE PROCEEDS VIA AN UNSTABLE ANHYDRIDE SPECIES FORMED
 CC BETWEEN THE CARBOXYLATE GROUPS OF THE ENZYME AND SUBSTRATE.
 CC -1- CATALYTIC ACTIVITY: SUCCINYL-COA + A 3-OXO ACID = SUCCINATE +
 CC A 3-OXO-ACYL-COA.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC -1- SIMILARITY: NO OTHER CO-TRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: M80534; AAA31019.1; -
 CC DR InterPro: IPR001618; -
 CC DR Pfam: PF01144; COA.trans. 2.
 CC DR PROSITE: PS01273; COA.TRANSF. 1;
 CC DR PROSITE: PS01274; COA.TRANSF. 2; 1.
 CC KW Mitochondrion; Transferase; Transit peptide.
 CC FT TRANSIT 1 39
 CC FT CHAIN 40 520
 CC FT DOMAIN 62 68
 CC FT ACT_SITE 344 344
 CC FT SEQUENCE 520 AA; 56407 MW; 3C9EA125FE78C222 CRC64;
 CC COA-BINDING (POTENTIAL).
 CC TRANSFERASE.

Query Match 71.1%; Score 32; DB 1; Length 520;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FTYTDAV 8
 DB 42 FTYTDAV 47

RESULT 6
 CSVR_ECOLI STANDARD; PRT; 301 AA.
 ID CSVR_ECOLI
 AC P43460;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE TRANSCRIPTIONAL ACTIVATOR CSVR.
 GN CSVR.
 OS Escherichia coli.

```

OG Plasmid.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EI0703 / SEROTYPE O167:H5;
RX MEDLINE=92120486; PubMed=1685133;
RA de Haan L.A., Willschaw G.A., van der Zeijst B.A., Gaastera W.;
RT "The nucleotide sequence of a regulatory gene present on a plasmid in
RL an enterotoxigenic Escherichia coli strain of serotype O167:H5."
CC FEMS Microbiol. Lett. 67:341-346(1991).
CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE FIMBRIAL GENE IN
CC ENTEROTOXIGENIC ESCHERICHIA COLI.
CC -1- SIMILARITY: BELONGS TO THE ARAC/XYLIS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60106; CAA42700.1; -
DR InterPro: IPR000005; -
DR Pfam: PF00165; HTH_ARAC. 1.
DR PRINTS: PR00032; HTHARAC.
DR PROSITE: PS00041; HTH_ARAC_FAMILY_2; 1.
DR PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
FT Transcription regulation; Activator; DNA-binding; Plasmid.
KW DNA_BIND 165 204 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 301 AA; 34887 MW; 7EFA307422DD256F CRC64;

Query Match
Best Local Similarity 68.9%; Score 31; DB 1; Length 301;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VTFYTDV 8
DB 161 IITFTDKV 168
:|||||
:|||||

RESULT 7
G3P_NEUCR STANDARD; PRT; 338 AA.
AC P54118; G92255;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (EC 1.2.1.12) (GAPDH)
DE (CLOCK-CONTROLLED PROTEIN 7).
GN GPD-1 OR CCG-7.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_Taxid=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC Saini M., Kinsey J.A.;
RT "Identification and cloning of the Neurospora crassa glyceraldehyde-3-
RL phosphate dehydrogenase gene, gpd-1."
RN Fungal Genet. Newsl. 44:47-49(1997).
RA Dunlap J.C., Shinohara M.L., Bell-Pedersen D., Ioros J.J.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) -> 1,3-DIPHOSPHATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

```

```

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U56397; AA00570.1; -
DR EMBL: U67457; AB93425.1; -
DR HSSP: P00357; IGPD.
DR InterPro: IPR000173; -
DR Pfam: PF00044; gpdh. 1.
DR PRINTS: PR00078; G3PDH_DGNASE.
DR PROSITE: PS00071; GAPDH. 1.
KW Glycolysis; Oxidoreductase; NAD.
FT BINDING 151 151
FT ACT_SITE 178 178
FT CONFLICT 78 79
FT CONFLICT 212 212 DA -> ER (IN REF. 2).
FT SEQUENCE 338 AA; 36193 MW; 62CBF5896D3B2F57 CRC64;

Query Match
Best Local Similarity 68.9%; Score 31; DB 1; Length 338;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VTFYTDV 7
DB 73 VKFYTDV 79
:|||||
:|||||

RESULT 8
SP5D_BACSU STANDARD; PRT; 645 AA.
AC Q03524;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE STAGE V SPOULATION PROTEIN D (SPOULATION SPECIFIC PENICILLIN-
DE BINDING PROTEIN).
GN SPOVD.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=94118264; PubMed=8289242;
RA Daniel R.A., Drake S., Buchanan C.E., Scholle R., Errington J.;
RT "The Bacillus subtilis SPOVD gene encodes a mother-cell-specific
RL penicillin-binding protein required for spore morphogenesis."
RN J. Mol. Biol. 235:209-220(1994).
RN [2]
RP SEQUENCE OF 1-69 FROM N.A.
RC STRAIN=168;
RX MEDLINE=94064553; PubMed=8244929;
RA Yanouri A., Daniel R.A., Errington J., Buchanan C.E.;
RT "Cloning and sequencing of the cell division gene pbpP, which encodes
RL penicillin-binding protein 2B in Bacillus subtilis."
RN J. Bacteriol. 175:7604-7616(1993).
RN [3]
RP SEQUENCE OF 595-645 FROM N.A.
RC STRAIN=168;
RX MEDLINE=93171879; PubMed=8436954;
RA Daniel R.A., Errington J.;

```

RT "DNA sequence of the mure-murd region of Bacillus subtilis 168.";
RL J. Gen. Microbiol. 139:361-370(1993).
RN [4]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN-168;
RA Daniel R.A., Williams A.M., Errington J.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: THIS IS A PENICILLIN-BINDING PROTEIN WITH AN UNKNOWN
CATALYTIC ACTIVITY, MAY HAVE A SPECIALIZED ROLE IN THE
MORPHOGENESIS OF SPORE CORTEX FORMATION IS DETERMINED PRIMARILY
BY THE MOTHER CELL.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (POTENTIAL).
CC -1- TISSUE SPECIFICITY: FOUND ONLY IN MOTHER CELLS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING SPOROLATION.
CC -1- SIMILARITY: TO VARIOUS OTHER PENICILLIN-BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 225865; CAB1085.1; -
DR EMBL; L09703; AAC36838.1; -
DR EMBL; Z15056; CAI7876.1; -
DR EMBL; Z68230; CAI92528.1; -
DR EMBL; Z69111; CAB13390.1; -
DR PIR; S23913; S23913.
DR PIR; A47691; A47691.
DR Subtilist; BG10222; spovd.
DR Interpro; IPR001460; -
DR Pfam; PF00905; Transpeptidase; 1.
DR Peptidoglycan synthesis; Cell wall; Sporulation; Membrane.
KW SEQUENCE 645 AA; 71261 MW; 0A0D896B6BDC2F4 CRC64;
SQ

Query Match 68.9%; Score 31; DB 1; Length 645;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFYTD 7
| | | | |
DB 191 VKFYTD 197

RESULT 9
METE_BACHD STANDARD; PRT; 756 AA.
AC 09KPEI;
DT 01-OCT-2000 (Rel. 40; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE 5-METHYLTERAHDROPTEROYLTRIGUTAMATE--HOMOCYSTEINE METHYLTRANSFERASE
(EC 2.1.1.14) (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME)
DE (CODALMIN-INDEPENDENT METHIONINE SYNTHASE).
GN METE OR BH0438.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RA MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuwara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- FUNCTION: CATALYZES THE TRANSFER OF A METHYL GROUP FROM 5-
METHYLTERAHDROPTERATE TO HOMOCYSTEINE RESULTING IN METHIONINE
FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 5-METHYLTERAHDROPTEROYLTR-L-GUTAMATE + L-
HOMOCYSTEINE = TETRAHDROPTEROYLTR-L-GUTAMATE + L-METHIONINE.
CC -1- COPACATOR: ZINC; BINDS ONE MOLE PER SUBUNIT (BY SIMILARITY).
CC -1- PATHWAY: TERMINAL STEP IN THE DE NOVO BIOSYNTHESIS OF METHIONINE.
CC -1- SIMILARITY: BELONGS TO THE VITAMIN-B12 INDEPENDENT METHIONINE
SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001508; BAB04157.1; -
DR TRANSFERASE; Methyltransferase; Methionine biosynthesis; Zinc; Repeat.
FT METAL 639 639 ZINC (BY SIMILARITY).
FT METAL 641 641 ZINC (BY SIMILARITY).
FT METAL 724 724 ZINC (BY SIMILARITY).
SQ SEQUENCE 756 AA; 85156 MW; 6F7FA9352A42AE6A CRC64;

Query Match 68.9%; Score 31; DB 1; Length 756;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VTFYTD 9
| | | | |
DB 243 VTFYERVS 251

RESULT 10
NID2_HUMAN STANDARD; PRT; 1375 AA.
AC 014112; 043710;
DT 15-DEC-1998 (Rel. 37; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE NIDDOGEN-2 PRECURSOR (NID-2) (OSTEONIDOGEN).
GN NID2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF N-TERMINUS.
RA MEDLINE=98406162; PubMed=9733643;
RA Kohleldt E., Sasaki T., Goehring W., Timpl R.;
RT "Nidogen-2: a new basement membrane protein with diverse binding
properties.";
RT J. Mol. Biol. 282:99-109(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cancellous bone;
RA Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.;
RT "The cloning and characterization of a cDNA for the novel bone matrix
protein: osteonidogen.";
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ohno I., Okubo K., Matsubara K.;
RT "Human osteonidogen gene: Intron-exon junctions and chromosomal
localization.";
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CELL ADHESION GLYCOPROTEIN WHICH IS WIDELY DISTRIBUTED
IN BASEMENT MEMBRANES. BINDS TO COLLAGENS I AND IV, TO PERLECAN
AND TO LAMININ 1. DOES NOT BIND FIBRININS. IT PROBABLY HAS A ROLE
IN CELL-EXTRACELLULAR MATRIX INTERACTIONS.
CC -1- TISSUE SPECIFICITY: HEART, PLACENTA AND BONE. LESS IN PANCREAS.

CC	KIDNEY AND SKELETAL MUSCLE.
CC	-1- PTM: HIGHLY N- AND O-GLYCOSYLATED.
CC	-1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC	-1- SIMILARITY: CONTAINS 2 THYROGLOBULIN TYPE-I DOMAINS.
CC	-1- SIMILARITY: CONTAINS 5 LDL-RECEPTOR YWTD DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AJ223500; CAA11418.1; -
DR	EMBL; D86425; BAA13087.1; -
DR	EMBL; AB009799; BAA24112.1; JOINED.
DR	EMBL; AB009778; BAA24112.1; JOINED.
DR	EMBL; AB009779; BAA24112.1; JOINED.
DR	EMBL; AB009780; BAA24112.1; JOINED.
DR	EMBL; AB009781; BAA24112.1; JOINED.
DR	EMBL; AB009782; BAA24112.1; JOINED.
DR	EMBL; AB009783; BAA24112.1; JOINED.
DR	EMBL; AB009784; BAA24112.1; JOINED.
DR	EMBL; AB009785; BAA24112.1; JOINED.
DR	EMBL; AB009786; BAA24112.1; JOINED.
DR	EMBL; AB009787; BAA24112.1; JOINED.
DR	EMBL; AB009788; BAA24112.1; JOINED.
DR	EMBL; AB009789; BAA24112.1; JOINED.
DR	EMBL; AB009790; BAA24112.1; JOINED.
DR	EMBL; AB009791; BAA24112.1; JOINED.
DR	EMBL; AB009792; BAA24112.1; JOINED.
DR	EMBL; AB009793; BAA24112.1; JOINED.
DR	EMBL; AB009794; BAA24112.1; JOINED.
DR	EMBL; AB009795; BAA24112.1; JOINED.
DR	EMBL; AB009796; BAA24112.1; JOINED.
DR	EMBL; AB009797; BAA24112.1; JOINED.
DR	EMBL; AB009798; BAA24112.1; JOINED.
DR	EMBL; AB009799; BAA24112.1; JOINED.
DR	EMBL; AB009800; BAA24112.1; JOINED.
DR	EMBL; AB009801; BAA24112.1; JOINED.
DR	EMBL; AB009802; BAA24112.1; JOINED.
DR	EMBL; AB009803; BAA24112.1; JOINED.
DR	EMBL; AB009804; BAA24112.1; JOINED.
DR	EMBL; AB009805; BAA24112.1; JOINED.
DR	EMBL; AB009806; BAA24112.1; JOINED.
DR	EMBL; AB009807; BAA24112.1; JOINED.
DR	EMBL; AB009808; BAA24112.1; JOINED.
DR	EMBL; AB009809; BAA24112.1; JOINED.
DR	EMBL; AB009810; BAA24112.1; JOINED.
DR	EMBL; AB009811; BAA24112.1; JOINED.
DR	EMBL; AB009812; BAA24112.1; JOINED.
DR	EMBL; AB009813; BAA24112.1; JOINED.
DR	EMBL; AB009814; BAA24112.1; JOINED.
DR	EMBL; AB009815; BAA24112.1; JOINED.
DR	EMBL; AB009816; BAA24112.1; JOINED.
DR	EMBL; AB009817; BAA24112.1; JOINED.
DR	EMBL; AB009818; BAA24112.1; JOINED.
DR	EMBL; AB009819; BAA24112.1; JOINED.
DR	EMBL; AB009820; BAA24112.1; JOINED.
DR	EMBL; AB009821; BAA24112.1; JOINED.
DR	EMBL; AB009822; BAA24112.1; JOINED.
DR	EMBL; AB009823; BAA24112.1; JOINED.
DR	EMBL; AB009824; BAA24112.1; JOINED.
DR	EMBL; AB009825; BAA24112.1; JOINED.
DR	EMBL; AB009826; BAA24112.1; JOINED.
DR	EMBL; AB009827; BAA24112.1; JOINED.
DR	EMBL; AB009828; BAA24112.1; JOINED.
DR	EMBL; AB009829; BAA24112.1; JOINED.
DR	EMBL; AB009830; BAA24112.1; JOINED.
DR	EMBL; AB009831; BAA24112.1; JOINED.
DR	EMBL; AB009832; BAA24112.1; JOINED.
DR	EMBL; AB009833; BAA24112.1; JOINED.
DR	EMBL; AB009834; BAA24112.1; JOINED.
DR	EMBL; AB009835; BAA24112.1; JOINED.
DR	EMBL; AB009836; BAA24112.1; JOINED.
DR	EMBL; AB009837; BAA24112.1; JOINED.
DR	EMBL; AB009838; BAA24112.1; JOINED.
DR	EMBL; AB009839; BAA24112.1; JOINED.
DR	EMBL; AB009840; BAA24112.1; JOINED.
DR	EMBL; AB009841; BAA24112.1; JOINED.
DR	EMBL; AB009842; BAA24112.1; JOINED.
DR	EMBL; AB009843; BAA24112.1; JOINED.
DR	EMBL; AB009844; BAA24112.1; JOINED.
DR	EMBL; AB009845; BAA24112.1; JOINED.
DR	EMBL; AB009846; BAA24112.1; JOINED.
DR	EMBL; AB009847; BAA24112.1; JOINED.
DR	EMBL; AB009848; BAA24112.1; JOINED.
DR	EMBL; AB009849; BAA24112.1; JOINED.
DR	EMBL; AB009850; BAA24112.1; JOINED.
DR	EMBL; AB009851; BAA24112.1; JOINED.
DR	EMBL; AB009852; BAA24112.1; JOINED.
DR	EMBL; AB009853; BAA24112.1; JOINED.
DR	EMBL; AB009854; BAA24112.1; JOINED.
DR	EMBL; AB009855; BAA24112.1; JOINED.
DR	EMBL; AB009856; BAA24112.1; JOINED.
DR	EMBL; AB009857; BAA24112.1; JOINED.
DR	EMBL; AB009858; BAA24112.1; JOINED.
DR	EMBL; AB009859; BAA24112.1; JOINED.
DR	EMBL; AB009860; BAA24112.1; JOINED.
DR	EMBL; AB009861; BAA24112.1; JOINED.
DR	EMBL; AB009862; BAA24112.1; JOINED.
DR	EMBL; AB009863; BAA24112.1; JOINED.
DR	EMBL; AB009864; BAA24112.1; JOINED.
DR	EMBL; AB009865; BAA24112.1; JOINED.
DR	EMBL; AB009866; BAA24112.1; JOINED.
DR	EMBL; AB009867; BAA24112.1; JOINED.
DR	EMBL; AB009868; BAA24112.1; JOINED.
DR	EMBL; AB009869; BAA24112.1; JOINED.
DR	EMBL; AB009870; BAA24112.1; JOINED.
DR	EMBL; AB009871

FT	DISULFID	852		867		BY SIMILARITY.
FT	DISULFID	859		877		BY SIMILARITY.
FT	DISULFID	879		890		BY SIMILARITY.
FT	DISULFID	896		907		BY SIMILARITY.
FT	DISULFID	901		916		BY SIMILARITY.
FT	DISULFID	918		929		BY SIMILARITY.
FT	CARBOHYD	417		417		N-LINKED (GLCNAC . . .) (PROBABLE).
FT	CARBOHYD	658		658		N-LINKED (GLCNAC . . .) (PROBABLE).
FT	CARBOHYD	693		693		N-LINKED (GLCNAC . . .) (PROBABLE).
FT	CARBOHYD	703		703		N-LINKED (GLCNAC . . .) (PROBABLE).
FT	CARBOHYD	1124		1124		N-LINKED (GLCNAC . . .) (PROBABLE).
FT	CONFLICT	44		44		G -> W (IN REF. 3).
FT	CONFLICT	54		172		D -> G (IN REF. 1).
SQ	SEQUENCE	1375 AA;	453	151394 MW;	414299D244205FBC CRC64;	
QY	1 VTFYPDAYS 9					
Db	1222 VLFTIDLVLN 1230					
RESULT 11	IDH2_YEAST					
AC	P28241:	STANDARD;		PRT;	369 AA.	
DT	01-DEC-1992 (Rel. 24, Created)					
DT	01-DEC-1992 (Rel. 24, Last sequence update)					
DT	15-JUL-1999 (Rel. 38, Last annotation update)					
DE	ISOCITRATE DEHYDROGENASE (NAD) SUBUNIT 2, MITOCHONDRIAL PRECURSOR					
DE	(EC 1.1.1.41) (ISOCITRIC DEHYDROGENASE) (NMD+-SPECIFIC ICDDH).					
GN	IDH2 OR YOR136W OR O3326 OR YOR3326W.					
OS	Saccharomyces cerevisiae (Baker's yeast).					
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;					
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.					
OX	NCBI_Taxid=4932;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=92042149; PubMed=1939242;					
RA	Cupp J.R., McAllister-Henn L.,					
RT	"NAD(+)-dependent isocitrate dehydrogenase. Cloning, nucleotide					
RT	sequence, and disruption of the IDH2 gene from Saccharomyces					
RL	cerevisiae."					
RL	J. Biol. Chem. 266:22199-22205(1991).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=5288C / FY1679;					
RX	MEDLINE=97060020; PubMed=8904341;					
RA	Wiemann S., Rechmann S., Benes V., Voss H., Schwager C., Vilek C.,					
RA	Stegemann J., Zimmermann J., Ertle H., Paces V., Ansoerge W.;					
RT	"Sequencing and analysis of 51 kb on the right arm of chromosome XV					
RT	from Saccharomyces cerevisiae reveals 30 open reading frames."					
RL	yeast 12:281-288(1996).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RA	MEDLINE=97344368; PubMed=9200815;					
RA	Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.,					
RA	Schwager C., Paces V., Sander C., Ansoerge W.;					
RT	"DNA sequencing and analysis of 130 kb from yeast chromosome XV."					
RT	yeast 13:655-672(1997).					
RL	[4]					
RN	SEQUENCE OF 16-26.					
RP	STRAIN-SG7;					
QC						

```

RX MEDLINE-90330530: PubMed-2198251;
RA Keys D.A., McAllister-Henn L.;
RT "Subunit structure, expression, and function of NAD(H)-specific
RL isocitrate dehydrogenase in Saccharomyces cerevisiae.";
RN J. Bacteriol. 172:4280-4287(1990).
RP RNA-BINDING, AND SEQUENCE OF 16-34.
RX MEDLINE-94089379: PubMed-7505425;
RA Elzinga S.D.J., Bednarek A.L., van Oosterum K., Dekker P.J.T.,
RA Grivell L.A.;
RT "Yeast mitochondrial NAD(+) dependent isocitrate dehydrogenase is an
RT RNA-binding protein.";
RL Nucleic Acids Res. 21:5328-5331(1993).
CC -1- FUNCTION: PERFORMS AN ESSENTIAL ROLE IN THE OXIDATIVE FUNCTION OF
CC THE CITRIC ACID CYCLE. ALSO BINDS RNA; SPECIFICALLY TO THE
CC 5'-UNTRANSLATED LEADERS OF MITOCHONDRIAL MRNAS.
CC -1- CATALYTIC ACTIVITY: ISOCITRATE + NAD(+) -> 2-OXOGLOUTARATE +
CC CO(2) + NADH.
CC -1- ENZYME REGULATION: ALLOSTERICALLY REGULATED BY SEVERAL COMPOUNDS
CC INCLUDING AMP, NAD+, AND CITRATE.
CC -1- SUBUNIT: OCTAMER OF TWO NONIDENTICAL SUBUNITS IDH1 AND IDH2.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
CC DEHYDROGENASES FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M74131; AAA34702.1; -
DR EMBL: X94335; CAA64054.1; -
DR EMBL: Z75043; CAA9335.1; -
DR EMBL: X90518; CAA62110.1; -
DR PIR: A39309; A39309.
DR HSSP: P00351; IOST.
DR SCD: S0005662; IDH2.
DR InterPro: IPR001804; -
DR Pfam: PF00180; Isodh; 1.
DR PROSITE: PS00470; IDH_LMDH; 1.
DR Oxidoreductase; NAD: Tricarboxylic acid cycle; Transit peptide;
KM Mitochondrion; Allosteric enzyme; RNA-binding.
FT TRANSIT 1 15
FT CHAIN 16 369 MITOCHONDRION.
FT ACT_SITE 113 113 ISOCITRATE DEHYDROGENASE [NAD] SUBUNIT 2.
FT CONFLICT 25 25 BINDING TO ISOCITRATE (BY SIMILARITY).
FT CONFLICT 25 25 R -> G (IN REF. 4).
SQ SEQUENCE 369 AA: 39739 MW: 3448399776CE373 CRC64;
-----
Query Match 66.7%; Score 30; DB 1; Length 369;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 YTDAYS 9
DB 250 YTDAYS 255
-----
RESULT 12
YL90_MYCTU STANDARD: PRT; 385 AA.
AC Q10383; OS3524;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOHETICAL 39.8 KDa PROTEIN RV2190C.
GN RV2190C OR MYC190.01C OR MT021.23C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.

```

```

OX NCBI_TaxID-1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE-98295987: PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AL021957; CAA17495.1; -
DR TubercuList; RV2190C; -
DR InterPro: IPR000064; -
DR Pfam: PF00877; NLPQ_P60; 1.
KW Hypothetical protein.
SQ SEQUENCE 385 AA: 39756 MW: 0F8F26577D0468DA CRC64;
-----
Query Match 66.7%; Score 30; DB 1; Length 385;
Best Local Similarity 71.4%; Pred. No. 80;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 YTDAYS 7
DB 340 YTDAYS 346
-----
RESULT 13
NM4M_RABIT STANDARD: PRT; 459 AA.
ID NM4M_RABIT
AC 079436;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4 (EC 1.6.5.3).
GN MTND4 OR ND4 OR NADH4.
OS Oryctolagus cuniculus (Rabbit).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID-9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98317530; PubMed-9653643;
RA Gissi C., Guilberg A., Arnason U.;
RT "The complete mitochondrial DNA sequence of the rabbit, Oryctolagus
RT cuniculus.";
RL Genomics 50:161-169(1998).
-----
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```


entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AF177936; AAD55772.1; -;
 DR EMBL: AF257109; AAF71745.1; -;
 DR HSSP: P03372; IHCP.
 DR InterPro: IPR000324; -;
 DR InterPro: IPR000536; -;
 DR InterPro: IPR001628; -;
 DR InterPro: IPR001723; -;
 DR Pfam: PF00104; hormone_rec; 1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PRO0047; STEROIDFINGER.
 DR PRINTS: PRO0350; VITAMINDR.
 DR PRINTS: PRO0398; STROHOMER.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation.
 FT DOMAIN 1 145 NUCLEAR RECEPTOR-TYPE.
 FT DNA_BIND 146 211
 FT ZN_FING 146 166 C4-TYPE.
 FT ZN_FING 182 206 C4-TYPE.
 FT DOMAIN 212 527 STEROID-BINDING.
 FT MOD_RES 84 84 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 485 485 PHOSPHORYLATION (BY SIMILARITY).
 FT VARSPLIC 316 324 FVELSLYDQ -> MKGNVLKEF (IN ISOFORM BETA-1).
 FT VARSPLIC 325 527 MISSING (IN ISOFORM BETA-1).
 SQ SEQUENCE 527 AA; 59120 MW; 9CD7A3E89497EDEA CRC64;

Query Match

Best Local Similarity 66.7%; Score 30; DB 1; Length 527;
 Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VTFYTDV 8
 ||||: ||
 DB 46 VTFYSPAV 53

Search completed: June 13, 2001, 14:30:37
 Job time: 527 sec

This Page Blank (uspto)

OC Bacteria; Firmicutes

OC Actinomycetaceae; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=68270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NRRL 2494;
 RA Altenbuchner J., Jansen D., Volf J.N.;
 RT "The spectinomycin resistance gene of the spectinomycin producer
 RT Streptomyces spectabilis.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF170704; AAD50451.1;
 DR INTERPRO: IPR000954; -;
 DR PFM: PF00202; aminotran_3; 1.
 FT NON_TER
 SQ SEQUENCE 98 AA; 10803 MW; 47E90E1AF4C752C8 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 98;
 Best Local Similarity 87.5%; Pred. No. 7;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 VFYTDAYS 9
 DB 86 VFYTDARS 93

RESULT 3
 ID 09KBP2 PRELIMINARY; PRT; 365 AA.
 AC 09KBP2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE BH1883 PROTEIN.
 GN BH1883.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RA Takami H., Nakasone K., Takaki Y.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP001513; BAB05602.1; -;
 SQ SEQUENCE 365 AA; 39312 MW; 69DF161C75A57692 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 365;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VFYTDAYS 9
 DB 138 VFYTDGDS 146

RESULT 4
 ID 09X6T5 PRELIMINARY; PRT; 442 AA.
 AC 09X6T5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE AMINOTRANSFERASE SPCS1.
 GN SPCS1.
 OS Streptomyces spectabilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=68270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC27741;
 RA Hyun C.G., Kim S.S., Suh J.W.;

RT "Isolation of two aminotransferase genes from Streptomyces spectabilis
 RT ATCC27741, a spectinomycin producer.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF145039; AAD28489.1; -;
 DR HSSP: P04181; 2CAN.
 DR INTERPRO: IPR000954; -;
 DR PFM: PF00202; aminotran_3; 1.
 DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; UNKNOWN_1.
 KW Transferase; Aminotransferase.
 SQ SEQUENCE 442 AA; 47574 MW; B2165F082EA4BB2E CRC64;

Query Match 75.6%; Score 34; DB 2; Length 442;
 Best Local Similarity 87.5%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 VFYTDAYS 9
 DB 86 VFYTDARS 93

RESULT 5
 ID 094831 PRELIMINARY; PRT; 968 AA.
 AC 094831;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE KIAA0726 PROTEIN.
 GN KIAA0726.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE-99087487; PubMed-9872452;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
 RA Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XI.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 5:277-286(1998).
 DR EMBL: AB018269; BAA34446.1; -;
 DR INTERPRO: IPR002126; -;
 DR PFM: PF00028; cadherin_2.
 SQ SEQUENCE 968 AA; 107033 MW; ECCB3E0214C50B8E CRC64;

Query Match 75.6%; Score 34; DB 4; Length 968;
 Best Local Similarity 66.7%; Pred. No. 96;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VFYTDAYS 9
 DB 471 VFYTDGDS 479

RESULT 6
 ID 09YKB9 PRELIMINARY; PRT; 1296 AA.
 AC 09YKB9; 09YKC0; 09Y023;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE BCDNA:LD23336 PROTEIN.
 GN BCDNA:LD23336.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

```

[1]
RN  SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP  STRAIN-BERKELEY;
RX  MEDLINE=20196006; PubMed=10731132;
RA  Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA  Abiri J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borovica D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA  Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA  Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Modarri C., Morris J., Mostrel A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA  Palazzolo M., Platten K.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA  Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster.";
RL  Science 287:2185-2195(2000).
RN  [2]
RP  SEQUENCE FROM N.A. (ISOFORM 3).
RC  STRAIN-BERKELEY;
RA  Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Broksstein P., Tsang G.,
RA  Aghayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA  Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Flise E.,
RA  Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA  Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacle J.M.,
RA  Park S., Sequeira A., Sehl H., Snir E., Svitskas R.R., Weinburg T.,
RA  Celniker S.E.;
RT  "Full length Drosophila melanogaster cDNA sequence.";
RL  Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC  -1- ALTERNATIVE PRODUCTS: THREE ISOFORMS: ISOFORM 1 (SHOWN HERE),
CC  ISOFORM 2 AND ISOFORM 3: ARE PRODUCED BY ALTERNATIVE SPLICING.
DR  EMBL: AE003634; AAF53156.1; -
DR  EMBL: AE003634; AAF53157.1; -
DR  EMBL: AE003634; -; NOT_ANNOTATED_CDS.
DR  EMBL: AF145682; AAD3657.1; -
DR  FLYBASE: FBgn0027505; BCdNA:LD23336.
KW  Alternative splicing.
FT  VARSPLIC 1 57 MISSING (IN ISOFORM 3).
FT  VARSPLIC 1 226 MISSING (IN ISOFORM 2).
FT  VARSPLIC 58 76 WTCVAVGLCSGWTFTYDS -> MDLRGSMPLYLMGTYLLHG
FT  L (IN ISOFORM 3).
SQ  SEQUENCE 1296 AA; 145953 MW; 9E7C13C1F4038D1A CRC64;

```

```

QY 1 VTFYTD 7
Db 70 VTFYTD 76
RESULT 7
ID 065728 PRELIMINARY; PRT; 154 AA.
AC 065728;
DT 01-AUG-1998 (TREMUREL. 07, Created)
DT 01-AUG-1998 (TREMUREL. 07, Last sequence update)
DT 01-JUN-2000 (TREMUREL. 14, Last annotation update)
DE HYPOPHYLLAL 16.9 KDA PROTEIN (FRAGMENT).
OS Cicer arietinum (chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Cicer.
OX NCBI_TaxID=3827;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, CASTELLANA; TISSUE=ETIOLATED EPICOTYL;
RA Labrador E., Esteban R., Dopico B.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: A7004959; CAA06244.1; -
DR INTERPRO: IPR002422; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 154 AA; 16918 MW; E5F992A1900D533D CRC64;

```

```

Query Match 73.3%; Score 33; DB 10; Length 154;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFYTD 9
Db 13 VTFYSAVS 21
RESULT 8
ID 09V224 PRELIMINARY; PRT; 472 AA.
AC 09V224;
DT 01-MAY-2000 (TREMUREL. 13, Created)
DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)
DT 01-MAY-2000 (TREMUREL. 13, Last annotation update)
DE ACETYL-COA DECARBOXYLASE/SYNTHASE BETA SUBUNIT.
GN CDHC.
OS Methanosarcina thermophila.
OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;
OC Methanosarcina.
OX NCBI_TaxID=2210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TM-1;
RA Gencic S., Grahame D.A.;
RT "Structure and function of acetyl-coA decarboxylase/synthase (ACDS)
RT protein subcomponents.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF173830; AAD51814.1; -
SQ SEQUENCE 472 AA; 52671 MW; C818616BB5C30A58 CRC64;

```

```

Query Match 73.3%; Score 33; DB 1; Length 472;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTFYTD 6
Db 148 VTFYTD 153
RESULT 9

```

09V010 ID 09V010 PRELIMINARY; PRT; 474 AA.
 AC 09V010;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE HYPOTHETICAL 52.6 KDA PROTEIN.
 GN PAB1825.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ORSAT;
 RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome structure and evolution";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ248285; CAB49723.1; -.
 KW Hypothetical protein.
 SO SEQUENCE 474 AA; 52571 MW; B50ED320F8CDC051 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 474;
 Best Local Similarity 75.0%; Pred. No. 69;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTTDAVS 9
 DB 281 TTTTDLIS 288

RESULT 10
 ID P91538 PRELIMINARY; PRT; 651 AA.
 AC P91538;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE COSMID ZC204.
 GN ZC204.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
 Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 Jones M., Kershaw J., Kristen J., Laister N., Latreille P.,
 Lighting J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RL Nature 366:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Wamsley P., Kramer J.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U80839; AAB37917.1; -.

DR INTERPRO: IPR001810; -;
 DR PFAM: PF00646; F-box; 2.
 SQ SEQUENCE 651 AA; 75446 MW; 05D3C67D51D7EAD2 CRC64;

Query Match 73.3%; Score 33; DB 5; Length 651;
 Best Local Similarity 85.7%; Pred. No. 99;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TTTTDAV 8
 DB 565 TTTTDTV 571

RESULT 11
 ID 09X5N2 PRELIMINARY; PRT; 689 AA.
 AC 09X5N2;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PROLYL ENDPEPTIDASE PRECURSOR PEP.
 GN PEP.
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 OC Myxococcales; Cytophasterineae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DZFL;
 RA Ueki T., Inouye S.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF127082; AAD31004.1; -.
 DR INTERPRO: IPR000379; -.
 DR INTERPRO: IPR001375; -.
 DR INTERPRO: IPR002470; -.
 DR INTERPRO: IPR002471; -.
 DR PFAM: PF00326; Peptidase_S9; 1.
 DR PRINTS: PR00862; PROLIGOPHASE.
 DR PROSITE: PS00708; PRO-ENDOPEP_SER; 1.
 SO SEQUENCE 689 AA; 76847 MW; 474FB9BB957D6DCB CRC64;

Query Match 73.3%; Score 33; DB 2; Length 689;
 Best Local Similarity 85.7%; Pred. No. 11e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 FYTDSVS 9
 DB 67 FYTDSVS 73

RESULT 12
 ID 09M2N3 PRELIMINARY; PRT; 695 AA.
 AC 09M2N3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HYPOTHETICAL 74.7 KDA PROTEIN.
 GN T2783.60.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mykatura G., Fatmann B., Dauner D., Sterr W., Holland R.,
 RA Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
 RA Salanoubat M.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ137079; CAB68120.1; -
 KW Hypothetical protein.
 SO SEQUENCE 695 AA; 74673 MW; C720E0ED69B41E98 CRC64;

Query Match 73.3%; Score 33; DB 10; Length 695;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TFEYTDVAVS 9
 Db 211 LSEYTDKVS 219

RESULT 13
 ID 075728 PRELIMINARY; PRT; 77 AA.
 AC 075728;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 GN VH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=INTESTINE;
 RA Fischer M., Kueppers R.;
 RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
 mutated VH region genes.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ009525; CAA08732.1; -
 DR INTERPRO: IPR003006; -
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1 77 1
 FT NON_TER 77 77 1
 SO SEQUENCE 77 AA; 8925 MW; 08EAEF6F1ECFBE6 CRC64;

Query Match 71.1%; Score 32; DB 4; Length 77;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VTFYTDVAV 8
 Db 26 VTFYPDVSV 33

RESULT 14
 ID 09PVZ1 PRELIMINARY; PRT; 312 AA.
 AC 09PVZ1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE C-JUN PROTEIN.
 GN C-JUN.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schuler-Metz A., Knoechel S., Knoechel W.;
 RT "Xvent-2 and c-Jun (AP-1) mediate auto-catalytic regulation of BMP-4
 expression in Xenopus embryos.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ243954; CAB51636.1; -
 DR HSSP: P05412; IPOS.
 DR INTERPRO: IPR001871; -
 DR INTERPRO: IPR002112; -
 DR PFAM: PF00170; bZIP; 1.
 DR PRINTS: PR00043; LEUZIPPRJUN.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 SO SEQUENCE 312 AA; 34363 MW; 3B9698703A4E3E1 CRC64;

Query Match 71.1%; Score 32; DB 13; Length 312;
 Best Local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 TFEYTDVAVS 9
 Db 8 TFEYDDALS 15

RESULT 15
 ID 09PVZ0 PRELIMINARY; PRT; 314 AA.
 AC 09PVZ0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE C-JUN PROTEIN.
 GN C-JUN.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schuler-Metz A., Knoechel S., Knoechel W.;
 RT "Xvent-2 and c-Jun (AP-1) mediate auto-catalytic regulation of BMP-4
 expression in Xenopus embryos.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ243955; CAB51637.1; -
 DR HSSP: P05412; IPOS.
 DR INTERPRO: IPR001871; -
 DR INTERPRO: IPR002112; -
 DR PFAM: PF00170; bZIP; 1.
 DR PRINTS: PR00043; LEUZIPPRJUN.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 SO SEQUENCE 314 AA; 34556 MW; 73396EA8CAC93B3 CRC64;

Query Match 71.1%; Score 32; DB 13; Length 314;
 Best Local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 TFEYTDVAVS 9
 Db 8 TFEYDDALS 15

Search completed: June 13, 2001, 14:29:44
 Job time: 545 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:07 ; Search time 78.71 Seconds

(without alignments)
2.197 Million cell updates/sec

Title: PCT-US01-05825A-26

Perfect score: 45

Sequence: 1 VFPTDAYS 9

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	9	2	US-08-859-931A-3	Sequence 3, Appl
2	31	68.9	119	4	US-08-983-607-22	Sequence 22, Appl
3	31	68.9	466	3	US-09-038-909-2	Sequence 2, Appl
4	30	66.7	116	3	US-09-184-658-48	Sequence 48, Appl
5	30	66.7	289	3	US-09-184-658-63	Sequence 63, Appl
6	30	66.7	291	5	PCT-US93-05000-6	Sequence 6, Appl
7	30	66.7	593	6	5523211-2	Sequence 6, Appl
8	30	66.7	595	6	5523211-3	Patent No. 5523211
9	30	66.7	613	2	US-08-622-740-6	Sequence 6, Appl
10	30	66.7	613	2	US-08-440-689-6	Sequence 6, Appl
11	30	66.7	614	2	US-08-622-740-8	Sequence 8, Appl
12	30	66.7	614	3	US-08-440-689-8	Sequence 8, Appl
13	30	66.7	615	4	US-09-042-426-10	Sequence 10, Appl
14	30	66.7	617	4	US-08-729-601A-2	Sequence 4, Appl
15	30	66.7	695	1	US-07-671-817A-5	Sequence 5, Appl
16	30	66.7	969	1	US-07-671-817A-4	Sequence 4, Appl
17	30	66.7	969	1	US-07-671-817A-6	Sequence 6, Appl
18	30	66.7	1155	1	US-08-349-867-19	Sequence 19, Appl
19	30	66.7	1155	1	US-08-349-867-33	Sequence 33, Appl
20	30	66.7	1155	1	US-08-239-476-13	Sequence 19, Appl
21	30	66.7	1155	1	US-08-446-486-2	Sequence 2, Appl
22	30	66.7	1155	1	US-07-951-715A-9	Sequence 9, Appl
23	30	66.7	1155	1	US-08-463-308-2	Sequence 2, Appl
24	30	66.7	1155	2	US-08-598-305A-19	Sequence 19, Appl
25	30	66.7	1155	2	US-08-598-305A-13	Sequence 33, Appl
26	30	66.7	1155	2	US-08-639-923A-19	Sequence 19, Appl
27	30	66.7	1155	2	US-08-459-448A-9	Sequence 9, Appl

28	30	66.7	1155	3	US-08-459-595A-9	Sequence 9, Appl
29	30	66.7	1155	3	US-09-021-203-2	Sequence 2, Appl
30	30	66.7	1155	3	US-08-459-504B-9	Sequence 9, Appl
31	30	66.7	1155	4	US-08-459-444-9	Sequence 9, Appl
32	30	66.7	1155	4	US-09-053-549-4	Sequence 4, Appl
33	30	66.7	1155	4	US-09-053-549-6	Sequence 4, Appl
34	30	66.7	1155	5	PCT-US95-05431-19	Sequence 19, Appl
35	30	66.7	1155	6	5254799-4	Patent No. 5254799
36	30	66.7	1165	1	US-08-239-476-33	Sequence 33, Appl
37	30	66.7	1165	2	US-08-639-923A-33	Sequence 33, Appl
38	30	66.7	1165	5	PCT-US95-05431-33	Sequence 33, Appl
39	30	66.7	1177	1	US-07-920-085-2	Sequence 8, Appl
40	30	66.7	1177	1	US-07-920-085-2	Sequence 2, Appl
41	30	66.7	1177	3	US-08-754-490-10	Sequence 10, Appl
42	30	66.7	1177	3	US-08-754-490-12	Sequence 12, Appl
43	30	66.7	1177	3	US-08-754-490-14	Sequence 14, Appl
44	30	66.7	1177	3	US-08-754-490-26	Sequence 26, Appl
45	30	66.7	1177	6	5169629-2	Patent No. 5169629

ALIGNMENTS

RESULT 1
US-08-859-931A-3
Sequence 3, Application US/08859931A
Patent No. 5945510
GENERAL INFORMATION:
APPLICANT: FASANO, Alessio
TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A
TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,931A
FILING DATE: 21 MAY 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6901
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-859-931A-3

Query Match 100.0%; Score 45; DB 2; Length 9;
Best local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
DB 1 VFPTDAYS 9
1 VFPTDAYS 9

RESULT 2
US-08-983-607-22
; Sequence 22, Application US/08983607
; Patent No. 6140470
; GENERAL INFORMATION:
; APPLICANT: Alan Garen
; APPLICANT: Xiaohong Cai
; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
; TITLE OF INVENTION: bodies
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Department of Molecular Biophysics
; ADDRESSEE: and Biochemistry, Yale University
; STREET: 266 Whitney Avenue
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: United States of America
; ZIP: 06520-8114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 MB diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processing
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,607
; FILING DATE: April 27, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/1996/01032
; FILING DATE: June 28, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Kinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: OCR-679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-773-8544
; TELEFAX: 203-773-1183
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; DESCRIPTION: polypeptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens (melanoma patient immu-
; ORGANISM: nized with autologous tumor cells)
; INDIVIDUAL ISOLATE: peritoneal blood lymphocytes
; IMMEDIATE SOURCE:
; LIBRARY: DM414 scfv antibodies obtained from
; LIBRARY: fuses fusion phage construct
; CLONE: G57
; FEATURE:
; NAME/KEY: heavy chain
; US-08-983-607-22

Query Match 68.9%; Score 31; DB 4; Length 119;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTYTDAV 8
|:|:|:|:
Db 58 TTYTDSV 64

RESULT 3
US-09-038-909-2
; Sequence 2, Application US/09038909

Patent No. 6083924
; GENERAL INFORMATION:
; APPLICANT: May, Earl W.
; APPLICANT: Earnshaw, David L
; APPLICANT: McDevitt, Damien
; TITLE OF INVENTION: dnab
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/038,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-038-909-2

Query Match 68.9%; Score 31; DB 3; Length 466;
Best Local Similarity 66.7%; Pred. No. 1,9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VTETDAVS 9
|:|:|:|:
Db 102 VQYTTDIVS 110

RESULT 4
US-09-184-658-48
; Sequence 48, Application US/09184658
; Patent No. 6030792
; GENERAL INFORMATION:
; APPLICANT: Olteneiss, Ivan G.
; APPLICANT: Mezes, Peter S.
; APPLICANT: Downs, James T.
; APPLICANT: Johnson, Kimberly S.
; TITLE OF INVENTION: Assays for Measurement of Protein Fragments in
; FILE REFERENCE: PC9946-A
; CURRENT APPLICATION NUMBER: US/09/184,658
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: 60/065,423
; EARLIER FILING DATE: 1997-11-13
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 116
; TYPE: PRT

ORGANISM: Mus musculus
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)..(116)
OTHER INFORMATION: Mature 5109 VH.
US-09-184-658-48

Query Match 66.7%; Score 30; DB 3; Length 116;
Best Local Similarity 62.5%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFTYDAV 8
:11111:
DB 57 LTFYADSV 64

RESULT 5
US-09-184-658-63
Sequence 63, Application US/09184658
Patent No. 6030792
GENERAL INFORMATION:
APPLICANT: Olteneess, Ivan G.
APPLICANT: Mezes, Peter S.
APPLICANT: Downs, James T.
APPLICANT: Johnson, Kimberly S.
TITLE OF INVENTION: Assays for Measurement of Protein Fragments In
FILE REFERENCE: PC9946-A
CURRENT APPLICATION NUMBER: US/09/184,658
CURRENT FILING DATE: 1998-11-02
EARLIER APPLICATION NUMBER: 60/065,423
EARLIER FILING DATE: 1997-11-13
NUMBER OF SEQ ID NOS: 69
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 63
LENGTH: 289
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 5109 VH - VL
OTHER INFORMATION: scfv.
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(22)
OTHER INFORMATION: pcantab6 signal peptide; Val at position 1 is most
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (23)..(138)
OTHER INFORMATION: 5109 VH domain.
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (139)..(154)
OTHER INFORMATION: 16 amino acid linker.
FEATURE:
NAME/KEY: SITE
LOCATION: (270)..(279)
OTHER INFORMATION: myc tag.
FEATURE:
NAME/KEY: SITE
LOCATION: (284)..(289)
OTHER INFORMATION: His tag.
US-09-184-658-63

Query Match 66.7%; Score 30; DB 3; Length 289;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFTYDAV 8
:11111:
DB 79 LTFYADSV 86

RESULT 6
PCT-US93-05000-6
Sequence 6, Application PC/TUS9305000
GENERAL INFORMATION:
APPLICANT: MITOXIX
TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/05000
FILING DATE: 19930525
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/888,178
FILING DATE: 26-MAY-1992.
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: CSHL91-02A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 616-861-9540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 291 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US93-05000-6

Query Match 66.7%; Score 30; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 YTDVAVS 9
:11111:
DB 127 YTDVAVS 132

RESULT 7
5523211-2
Patent No. 5523211
APPLICANT: PUSZTAI-CAREY, MARIANNE; CAREY, PAUL R.; LESSARD,
TIMOTHY; YAGICH, MAKORO
TITLE OF INVENTION: IDENTIFICATION, QUANTITATION AND
PURIFICATION OF INSECTICIDAL PROTEINS FROM BACILLUS THURINGIENSIS
NUMBER OF SEQUENCES: 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/277,076
FILING DATE: 19-JUL-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 102,491
FILING DATE: 05-AUG-1993
APPLICATION NUMBER: 836,967
FILING DATE: 19-FEB-1992

APPLICATION NUMBER: 493,453
FILING DATE: 14-MAR-1990
SEQ ID NO: 2
LENGTH: 593
5523211-2

Query Match
Best Local Similarity 66.7%; Score 30; DB 6; Length 593;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 VTFYTD 7
: 11111
DB 274 ITTYTDA 280

RESULT 8
5523211-3
PATENT NO. 5523211
APPLICANT: PUSZTAI-CAREY, MARIANNE; CAREY, PAUL R.; LESSARD,
TIMOTHY; YAGUCHI, MAKOTO
TITLE OF INVENTION: IDENTIFICATION, QUANTITATION AND
PURIFICATION OF INSECTICIDAL PROTEINS FROM BACILLUS THURINGIENSIS
NUMBER OF SEQUENCES: 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/277,076
FILING DATE: 19-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 102,491
FILING DATE: 05-AUG-1993
APPLICATION NUMBER: 836,967
FILING DATE: 19-FEB-1992
APPLICATION NUMBER: 493,453
FILING DATE: 14-MAR-1990
SEQ ID NO: 3
LENGTH: 595
5523211-3

Query Match
Best Local Similarity 66.7%; Score 30; DB 6; Length 595;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 VTFYTD 7
: 11111
DB 275 ITTYTDA 281

RESULT 9
US-08-622-740-6
Sequence 6, Application US/08622740
Patent No. 5990390
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
TITLE OF INVENTION: Production of Stably Transformed, Fertile Monocot Plants
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/622,740
FILING DATE: 27-MARCH-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 950.013US4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 613 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-622-740-6

Query Match
Best Local Similarity 66.7%; Score 30; DB 2; Length 613;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 VTFYTD 7
: 11111
DB 303 ITTYTDA 309

RESULT 10
US-08-440-689-6
Sequence 6, Application US/08440689
Patent No. 6025545
GENERAL INFORMATION:
APPLICANT: Lundquist, Ronald C.
APPLICANT: Walters, David A.
APPLICANT: Kirihara, Julie A.
TITLE OF INVENTION: Methods and Compositions for the
TITLE OF INVENTION: Production of Stably Transformed, Fertile Monocot Plants
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,689
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 950.013US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 613 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-689-6

Query Match
66.7%; Score 30; DB 3; Length 613;

Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VFFYTD 7
: 1 1111
Db 303 ITTYTDA 309

RESULT 11

US-08-622-740-8
; Sequence 8, Application US/08622740
; Patent No. 5990390
; GENERAL INFORMATION:
; APPLICANT: Lundquist, Ronald C.
; APPLICANT: Walters, David A.
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Production of Stably Transformed, Fertile Monocot Plants
; TITLE OF INVENTION: and Cells Thereof
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,740
; FILING DATE: 27-MARCH-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 614 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-622-740-8

Query Match 66.7%; Score 30; DB 2; Length 614;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VFFYTD 7
: 1 1111
Db 303 ITTYTDA 309

RESULT 12

US-08-440-689-8
; Sequence 8, Application US/08440689
; Patent No. 6025545
; GENERAL INFORMATION:
; APPLICANT: Lundquist, Ronald C.
; APPLICANT: Walters, David A.
; APPLICANT: Kirlhara, Julie A.
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Production of Stably Transformed, Fertile Monocot Plants
; TITLE OF INVENTION: and Cells Thereof
; NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,689
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 614 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-689-8

Query Match 66.7%; Score 30; DB 3; Length 614;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VFFYTD 7
: 1 1111
Db 303 ITTYTDA 309

RESULT 13

US-09-042-426-10
; Sequence 10, Application US/09042426
; Patent No. 6114608
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6114608artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,426
; FILING DATE: March 13, 1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 615 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
IMMEDIATE SOURCE:
CLONE: Bt protein
US-09-042-426-10

Query Match 66.7%; Score 30; DB 4; Length 615;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFTDA 7
: 1 1 1 1 1
Db 303 ITTYTDA 309

RESULT 14
US-08-729-601A-2
Sequence 2, Application US/08729601A
Patent No. 6166302
GENERAL INFORMATION:
APPLICANT: Merlo, Donald J.
APPLICANT: Folkerts, Otto
TITLE OF INVENTION: Modified Bacillus Thuringiensis Gene for
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle St.
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,601A
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Krueger, James P.
REGISTRATION NUMBER: 35,234
REFERENCE/DOCKET NUMBER: 60089
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-372-7842
TELEFAX: 312-372-7848
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-729-601A-2

Query Match 66.7%; Score 30; DB 4; Length 617;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFTDA 7
: 1 1 1 1 1
Db 303 ITTYTDA 309

RESULT 15
US-07-671-817A-5
Sequence 5, Application US/07671817A
Patent No. 5424409
GENERAL INFORMATION:
APPLICANT: Ely, Susan
APPLICANT: Tippet, Janet M
TITLE OF INVENTION: DNA constructs
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby and Cushman
STREET: Eleventh floor, 1615 L Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-3601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/671,817A
FILING DATE: 19910401
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8823068.5
FILING DATE: 30-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB89/01157
FILING DATE: 29-SEP-1989
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 695 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-671-817A-5

Query Match 66.7%; Score 30; DB 1; Length 695;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTFTDA 7
: 1 1 1 1 1
Db 303 ITTYTDA 309

Search completed: June 13, 2001, 14:27:07
Job time: 628 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:46 ; Search time 150.28 Seconds
(without alignments)
7.608 Million cell updates/sec

Title: PCT-US01-05825A-27

perfect score: 93

Sequence: 1 MQLAESGVLVQPGXSDDL 20

Scoring table:

BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

1:	/SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	86.0	20	W94489	Human foetal intest
2	80	86.0	20	W94489	Human foetal intest
3	66	71.0	98	21 W94489	Anti-HiLi2 antibody
4	66	71.0	118	16 R66331	Human immunoglobulin
5	65	69.9	124	19 W80816	Amino acid sequenc
6	64	68.8	20	21 Y84660	N-terminal sequenc
7	64	68.8	98	15 R54816	SPA-reactive VH re
8	63	67.7	134	15 R52807	Humanised murine B
9	63	67.7	253	20 W90225	Anti-B7.1 monospec
10	63	67.7	258	20 W90221	Anti-B7.1 monospec
11	63	67.7	556	20 W90218	Bispecific tetra

12	63	67.7	580	20	W90217	Bispecific tetra
13	62	66.7	20	19	W68169	Human IgG heavy ch
14	62	66.7	20	20	W94487	Human adult heart
15	62	66.7	20	21	Y79130	Human adult heart
16	62	66.7	25	15	R52775	Murine BRE-3 immu
17	62	66.7	25	15	R52793	Murine BRE-3 immu
18	62	66.7	30	17	R87049	Human group III he
19	62	66.7	84	21	Y64925	Human 5' EST relat
20	62	66.7	85	21	Y64735	Human 5' EST relat
21	62	66.7	96	21	Y64844	Human 5' EST relat
22	62	66.7	97	21	B40112	Anti-HiLi2 antibod
23	62	66.7	98	14	R34279	Human TNF binding
24	62	66.7	98	15	R52066	Heavy chain variab
25	62	66.7	98	16	R72074	DP54 VH region. H
26	62	66.7	98	19	W59614	Anti-RSV F protein
27	62	66.7	98	21	B40072	Anti-HiLi2 antibod
28	62	66.7	98	21	B40084	Anti-HiLi2 antibod
29	62	66.7	98	21	B40085	Anti-HiLi2 antibod
30	62	66.7	98	21	B40086	Anti-HiLi2 antibod
31	62	66.7	98	21	B40087	Anti-HiLi2 antibod
32	62	66.7	98	21	B40094	Anti-HiLi2 antibod
33	62	66.7	98	21	B40095	Anti-HiLi2 antibod
34	62	66.7	98	21	B40096	Anti-HiLi2 antibod
35	62	66.7	98	21	B40107	Anti-HiLi2 antibod
36	62	66.7	98	21	B40108	Anti-HiLi2 antibod
37	62	66.7	98	21	B40113	Anti-HiLi2 antibod
38	62	66.7	98	21	B40130	Anti-HiLi2 antibod
39	62	66.7	98	21	B40134	Anti-HiLi2 antibod
40	62	66.7	98	21	B40136	Anti-HiLi2 antibod
41	62	66.7	98	21	B40137	Anti-HiLi2 antibod
42	62	66.7	98	21	B40138	Anti-HiLi2 antibod
43	62	66.7	98	21	B40139	Anti-HiLi2 antibod
44	62	66.7	98	21	B40140	Anti-HiLi2 antibod
45	62	66.7	98	21	Y56660	Partial peptide fr

ALIGNMENTS

RESULT 1	
ID	W94489
W94489	standard; peptide: 20 AA.
AC	W94489:
XX	
XX	21-APR-1999 (first entry)
XX	
XX	Human foetal intestine zonulin N-terminal peptide.
DE	
XX	
XX	Zonulin: mammalian tight junction; zonula occludens toxin; ZOT;
KW	Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody;
KW	intestinal mucosa; nasal mucosa; blood brain barrier.
XX	
OS	Homo sapiens.
XX	
XX	
FT	Key
FT	Misc-difference 16
FT	/note= "unspecified"
XX	
XX	W09852415-A1.
XX	
XX	26-NOV-1998.
XX	
XX	28-APR-1998; 98WO-US07636.
XX	
XX	21-MAY-1997; 97US-0859931.
XX	
XX	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
XX	SPASANO A;
XX	
XX	Fasano A;
XX	
XX	WPI; 1999-070123/06.
XX	

Claim 2; Page 45; 64pp; English.

Sequence 20 AA;

Db 1 mlqkaesgvlvqpgxsnrl 20

AC Y79133;

DT 05-JUN-2000 (first entry)

Human foetal intestine zonulin N-terminal sequence
De
XX

gastrointestinal inflammation; therapy

OS Homo sapiens.

EH	Key	Location/Qualifiers
ET	Misc-differences	16

f1 /note= "unidentified residue
xx

PN MO200007609-A1.
XY

PD 17-FEB-2000

PF 28-JUL-1999; 99WO-US16683

PR 03-AUG-1998; 98US-0127815.

PA (UYMA-) UNIV MARYLAND BALTIMORE
XX

PI Fasano A;

DR WPI; 2000-205565/18.

PS Example 3; Fig 6; 69pp; English.
XY

CC barrier.

50 Sequence 20 AA

1 mlqkaesgv1vqpgxsnr1 20

ID	B40074	standard; Protein; 98 AA
----	--------	--------------------------

AC B40074;
YY

03-FEB-2001 (first entry)

XXXXXX amino acid sequence SEQ ID NO 600

multiple sclerosis; rheumatoid arthritis

XX
XX

nullo sapientis.

XX	1
XX	2
XX	3
XX	4
XX	5
XX	6
XX	7
XX	8
XX	9
XX	10
XX	11
XX	12
XX	13
XX	14
XX	15
XX	16
XX	17
XX	18
XX	19
XX	20
XX	21
XX	22
XX	23
XX	24
XX	25
XX	26
XX	27
XX	28
XX	29
XX	30
XX	31
XX	32
XX	33
XX	34
XX	35
XX	36
XX	37
XX	38
XX	39
XX	40
XX	41
XX	42
XX	43
XX	44
XX	45
XX	46
XX	47
XX	48
XX	49
XX	50
XX	51
XX	52
XX	53
XX	54
XX	55
XX	56
XX	57
XX	58
XX	59
XX	60
XX	61
XX	62
XX	63
XX	64
XX	65
XX	66
XX	67
XX	68
XX	69
XX	70
XX	71
XX	72
XX	73
XX	74
XX	75
XX	76
XX	77
XX	78
XX	79
XX	80
XX	81
XX	82
XX	83
XX	84
XX	85
XX	86
XX	87
XX	88
XX	89
XX	90
XX	91
XX	92
XX	93
XX	94
XX	95
XX	96
XX	97
XX	98
XX	99
XX	100

9
 8
 7
 6
 5
 4
 3
 2
 1

XX
DB 3E-MAD-1000 0012 01000000

PA (BADI) BASE AC

(GENI) GENELICS INST INC.

PI Kaymakçalan Z, Labkovsky B, Sakorafas P, Friedrich S, Myles A;

Derbyshire EJ, Carmen S, Smith S, Holtet TL, Du Fou SL;

DR
WPI; 2000-638250/61.

New human antibody specific for human interleukin-12 (IL-12) used to treat disorders characterized by aberrant IL-12 overexpression

increase and multiple sclerosis -

Claim 15; Page 121; 37/pp; English.

CC Interleukin-12 (IL-12). The invention also includes antigen binding
 CC portions that bind to IL-12. Sequences B39485-B39516 represent human
 CC anti-IL-12 antibody heavy and light chain complementarily determining
 CC region (CDR) amino acid sequences, and also includes variable region
 CC amino acid sequences. Other variable region amino acid sequences are
 CC given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
 CC represent anti-IL-12 CDR3 related amino acid sequences, B39772-B40063
 CC represent other CDR sequences. Light chain CDR3 consensus sequences are
 CC given in B40064-B40067. Primers used in the identification and
 CC construction of the antibodies of the invention are given in
 CC C61062-C61071. The antibody of the invention is a neutralising antibody
 CC and has antirheumatic; antiallergic; antisclerotic; antiinflammatory;
 CC neuroprotective; antipapillary; antistomatitis; candida; antiparasitic;
 CC antibacterial and immunosuppressive activity. The antibodies or
 CC antigen-binding fragments are useful in the treatment of disorders
 CC associated with detrimental release of human IL-12, especially Crohn's
 CC disease, multiple sclerosis and rheumatoid arthritis. They can also be
 CC used in the manufacture of a pharmaceutical composition to treat human
 CC IL-12 disorders.

CC Sequence 98 AA:

Query Match 71.0%; Score 66; DB 21; Length 98;
 Best Local Similarity 73.7%; Pred. No. 0.0036;
 Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LQAESGVLVOPGXSDRL 20
 :||| |||||:||||| |||
 Db 2 VQLVESGVLVVPQPSLRL 20

RESULT 4
 R66331 ID R66331 standard; Protein: 118 AA.

AC R66331;
 XX
 XX 04-AUG-1995 (first entry)

DE Human Immunoglobulin variable heavy chain #37.

XX Primer: PCR; amplify: human; Immunoglobulin; variable; heavy chain;
 KW cosmid; placenta; vector; pJB81; E.coli; mammalian.

XX Homo sapiens.

XX MO9426895-A.

XX 24-NOV-1994.

XX 10-MAY-1993; 93WO-JP00603.

XX 10-MAY-1993; 93WO-JP00603.

XX (NLSB) JAPAN TOBACCO INC.

XX Honjo T, Matsuda F;

XX WPI: 1995-006791/01.

XX N-PSDB: Q78981.

PT DNA fragment comprising human immunoglobulin Vh genes - for the
 PT production of human immunoglobulin in mammalian hosts

XX Claim 50; Page 85-87; 130pp; Japanese.

CC Protein sequences (R66295-51) are novel human immunoglobulin heavy chain
 CC sequences encoded by novel isolated genes. The genes (Q78939-79002) were
 CC isolated and cloned from a series of cosmid constructs: Y202; Y103; Y21;
 CC Y6; Y24; 3-31; M84; M18 and M31, by PCR amplification using primers
 CC Q78917-38. The genes are subdivided into 5 families of Vh genes. The
 CC fragments cover a region of 800 kb. The DNA fragments were isolated from

CC high molecular weight DNA from human placenta. The DNA was partially
 CC digested with RsaI restriction enzyme. The fragments were separated by
 CC gel electrophoresis and 35-45 kb fractions were collected. The fragments
 CC were ligated with ClaI-digested cosmid vector pJB81. The ligation
 CC products were in vitro packed and infected into E.coli 490A. The
 CC fragments were then subcloned by colony hybridisation. The Vh genes and
 CC the DNA fragments encoding them are useful in producing human
 CC immunoglobulin in mammalian hosts.

CC Sequence 118 AA:

Query Match 71.0%; Score 66; DB 16; Length 118;
 Best Local Similarity 73.7%; Pred. No. 0.0043;
 Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LQAESGVLVOPGXSDRL 20
 :||| |||||:||||| |||
 Db 21 VQLVESGVLVVPQPSLRL 39

RESULT 5
 W80816 ID W80816 standard; Protein: 124 AA.

AC W80816;

DE 16-FEB-1999 (first entry)

XX Amino acid sequence of human D7.2 heavy chain variable region.

XX Human; D7.2 heavy chain variable region; receptor; antigen;

XX tumour; auto-immune disease; graft rejection; allergy;

XX inflammatory disease; endocrine disease; degenerative disease.

XX Homo sapiens.

XX WO9846645-A2.

XX 22-OCT-1998.

XX 14-APR-1998; 98WO-EP02180.

XX 14-APR-1997; 97EP-0106109.

XX (KUFE/) KUFE P.

XX (RAUM/) RAUM T.

XX Kufer P, Raum T;

XX WPI: 1998-594564/50.

XX N-PSDB: V68538.

XX Production of anti-human antigen receptors - by selecting a
 XX combination of functionally rearranged Vh and VL immunoglobulin
 XX chains expressed from a recombinant vector

XX Claim 9; Fig 8; 84pp; English.

CC This is the amino acid sequence of the human D7.2 heavy chain
 CC variable region, used in the method of the invention. for providing
 CC receptors that can be used for targeting antigens in humans without
 CC being immunogenic themselves. Such receptors can be used for treating
 CC diseases such as tumours or auto-immune diseases, graft rejection
 CC after transplantation, infectious diseases by targeting cellular
 CC receptors as well as allergic, inflammatory, endocrine and
 CC degenerative diseases by targeting key molecules involved in the
 CC pathological process.

CC Sequence 124 AA:

Query Match 69.9%; Score 65; DB 19; Length 124;

OY 2 LQLAESGVLVOPGXSDRL 20
 ||| ||||| ||||| |||
 DB 2 lqllesgg9glvqpg9srl 20

RESULT 8
 R52807
 ID R52807 standard; Protein: 134 AA.
 XX
 AC R52807;
 XX
 DT 24-JAN-1995 (first entry)
 XX
 DE Humanised murine B73 immunoglobulin heavy chain V-region.
 XX
 KW Immunoglobulin variable domain; primer: polymerase chain reaction;
 KM chimeric antibody; human mammary fat globule; human breast carcinoma;
 KM murine anti-HMEG monoclonal antibody B73; humanised analogue.
 XX
 OS Chimeric Mus musculus.
 OS Chimeric Homo sapiens.
 XX

Key Location/Qualifiers
 FH 1..19
 FT /label= leader
 FT Region 20..49
 FT /label= FR1
 FT /note= "Claim 4"
 FT 50..54
 FT /label= CDRI
 FT /note= "Claim 4"
 FT 55..68
 FT /label= FR2
 FT /note= "Claim 4"
 FT 69..87
 FT /label= CDR2
 FT /note= "Claim 4"
 FT 88..119
 FT /label= FR3
 FT /note= "Claim 4"
 FT 120..123
 FT /label= CDR3
 FT /note= "Claim 4"
 FT 124..134
 FT /label= FR4
 FT /note= "Claim 4"

WO9411509-A.
 XX
 PD 26-MAY-1994.
 XX
 PF 16-NOV-1993; 93WO-US11445.
 XX
 PR 16-NOV-1992; 92US-0977696.
 PR 30-SEP-1993; 93US-0129930.
 PR 08-OCT-1993; 93US-0134346.
 XX
 PA (CANC-) CANCER RES FUND CONTRA COSTA.
 XX
 PI Ceriani RL, Do Couto FJR, Padlan EA, Peterson JA;
 XX
 DR WPI, 1994-183510/22.
 DR N-PSDB: Q62791.
 XX
 XX New analogue peptide(s) comprising antibody variable regions -
 PT used to develop prods. for use in the detection, diagnosis,
 PT therapy and prevention of neoplasms
 XX
 PS Claim 4; Page 84; 109pp; English.
 XX
 CC The different regions of the mature polypeptide R52807 indicated in
 CC the Features Table are preferred partial sequences for antineoplastic

CC analogue peptides. The peptides selectively bind to an antigen on the
 CC surface or in the cytoplasm of a neoplastic cell. They comprise at
 CC least one CDR and at least one V region of L- or H-chains from an
 CC anti-human mammary fat globule antigen antibody (i.e. murine B73
 CC antibody) but where 1 to 46 amino acids in the FRs are substituted by
 CC their homologues from human antibodies (humanised). The humanised
 CC antibody is useful for carcinoma therapy and diagnosis and for in
 CC vivo imaging of neoplastic cells. It is also of use in inhibiting the
 CC growth of a primary or metastasised neoplasm.
 XX

SQ Sequence 134 AA;
 XX

Query Match 67.7%; Score 63; DB 15; Length 134;
 Best Local Similarity 73.7%; Pred. No. 0.015;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSDRL 20
 :|| ||||| ||||| |||
 DB 21 vlqvesgg9glvqpg9saml 39

RESULT 9
 W90225
 ID W90225 standard; Protein: 253 AA.
 XX
 AC W90225;
 XX
 DT 10-MAY-1999 (first entry)
 XX
 DE Anti-B7.1 monospecific triabody B7-24.
 XX
 KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KM CD86; T cell activation; inhibitor; graft versus host disease;
 KM transplant rejection; allograft rejection; autoimmune disease;
 KM allergy; therapy; human; triabody; antibody; B7-24.
 XX
 OS Chimeric - Mus sp.
 OS Chimeric - synthetic.
 XX

Key Location/Qualifiers
 FH 1..24
 FT /note= "pe1b signal peptide"
 FT Region 25..138
 FT /note= "anti B7.1 MAb VH region"
 FT 139..247
 FT /note= "anti B7.1 MAb VL region"
 FT 248..253
 FT /note= "His6 tag"

WO9858965-A2.
 XX
 PD 30-DEC-1998.
 XX
 PF 22-JUN-1998; 98WO-EP03791.
 XX
 PR 20-JUN-1997; 97EP-0870092.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Bosman A, Buyse M, Lorre K, Sablon E;
 XX
 DR WPI, 1999-105615/09.
 DR N-PSDB: X01659.
 XX
 XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 PT immune diseases including allograft rejection
 XX
 PS Example 7.3; Fig 32; 182pp; English.
 XX
 CC This polypeptide comprises a B7-24 monospecific triabody composed
 CC of the VH region of anti-B7.1 monoclonal antibody (MAb) B7-24
 CC joined to the VL region of B7-24. A triabody is a mono- a bi- or

CC a trispecific molecule recognising simultaneously e.g. two B7.2
 CC and one B7.1 molecules. It has a rigid structure that prevents
 CC simultaneous binding to the 3 targets. Each antigen binding site
 CC is formed by pairing of one VH and one VL domain from the same or
 CC from two different polypeptides. The invention relates to novel
 CC molecules, including triabodies, which can cross-link and/or
 CC cross-react with the costimulatory molecules B7.1 and B7.2 expressed
 CC on professional antigen-presenting cells, leading to the inhibition
 CC of antigen-specific T cell activation. Methods are provided for
 CC the production of such B7-binding molecules, and for their use in
 CC the treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).

XX Sequence 253 AA;

Query Match 67.7%; Score 63; DB 20; Length 253;
 Best Local Similarity 73.7%; Pred. No. 0.029;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGGLVDPGXSDRL 20
 Db 24 vqlqesggivpypgsrrl 42

RESULT 10
 W90221
 ID W90221 standard; Protein; 258 AA.

XX W90221;

DT 10-MAY-1999 (first entry)

DE Anti-B7.1 monospecific diabody B7-24.

XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KM CD86; T cell activation; inhibitor; graft versus host disease;
 KM transplant rejection; allograft rejection; autoimmune disease;
 KM allergy; therapy; human; diabody; antibody; B7-24.

XX Chimeric - Mus sp.
 OS Chimeric - synthetic.

XX Key Location/Qualifiers

FT Peptide 1..24
 FT /note="pe1b signal peptide"

FT Region 25..138
 FT /note="anti B7.1 Mab VH region"

FT Peptide 139..143
 FT /note="G4S flexible linker"

FT Region 144..252
 FT /note="anti B7.1 Mab VL region"

FT Peptide 253..258
 FT /note="His6 tag"

FT Peptide

PN W09858965-A2.

XX 30-DEC-1998.

XX 22-JUN-1998; 98MO-EP03791.

XX 20-JUN-1997; 97EP-0870092.

XX (INNO-) INNOGENETICS NV.

XX Bosman A, Buyse M, Lorre K, Sablon E;

XX WPI; 1999-105615/09.

XX DR N-PSDB; X01655.

PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 PT immune diseases including allograft rejection

XX Example 7.2; Fig 24; 182pp; English.

PS This polypeptide comprises a B7-24 monospecific diabody composed
 CC of the VH region of anti-B7.1 monoclonal antibody (Mab) B7-24
 CC joined via a short, flexible linker to the VL region of B7-24.
 CC Mono- or bispecific bivalent molecules are generated by shortening
 CC the flexible linker sequence between the VH and VL of the anti-B7.1
 CC scFv B7-24, the anti-B7.2 scFv 1G10 and the scFv molecule with
 CC dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and
 CC for bispecific molecules by cross-pairing the VH and VL domains
 CC from the 2 scFvs with different antigen recognition (B7.1/B7.2 and
 CC B7.12/B7.12). The invention relates to novel molecules, including
 CC diabodies, which can cross-link and/or cross-react with the
 CC costimulatory molecules B7.1 and B7.2 expressed on professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. Methods are provided for the
 CC production of such B7-binding molecules, and for their use in the
 CC treatment or prevention of diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).

XX Sequence 258 AA;

Query Match 67.7%; Score 63; DB 20; Length 258;
 Best Local Similarity 73.7%; Pred. No. 0.03;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGGLVDPGXSDRL 20
 Db 24 vqlqesggivpypgsrrl 42

RESULT 11

W90218
 ID W90218 standard; Protein; 556 AA.

XX W90218;

DT 10-MAY-1999 (first entry)

DE Bispecific tetravalent antibody B17AD1G10-B7-24H6.

XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 KM CD86; T cell activation; inhibitor; graft versus host disease;
 KM transplant rejection; allograft rejection; autoimmune disease;
 KM allergy; therapy; human; bispecific tetravalent antibody; B17Ab;
 KM B17AD1G10-B7-24H6.

XX Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Chimeric - synthetic.

XX Key Location/Qualifiers

FT Region 1..120
 FT /note="VH region anti B7.2 Mab"

FT Peptide 121..135
 FT /note="(G4S3) flexible linker"

FT Region 136..248
 FT /note="VL region anti B7.2 Mab"

FT Region 249..259
 FT /note="human IgG3 hinge region"

FT Domain 260..285
 FT /note="helix-turn-helix dimerisation domain"

FT Domain 286..305
 FT /note="human IgG3 hinge domain"

FT Region 306..426
 FT /note="VH region anti B7.1 Mab"

FT Peptide 427..441
 FT /note="(G4S3)flexible linker"

FT Region 442..550
 FT /note="VL region anti B7.1 Mab"

FT Peptide 551..556
 FT /note="His6 tag"
 XX
 XX W09858965-A2.
 XX
 XX 30-DEC-1998.
 PD
 XX
 XX 22-JUN-1998: 98WO-EP03791.
 XX
 XX 20-JUN-1997: 97EP-0870092.
 PR
 XX (INNO-) INNOGENETICS NV.
 PA
 XX Bosman A, Buyse M, Lorre K, Sablon E;
 PI
 XX WPI: 1999-105615/09.
 DR
 XX N-PSDB: X01652.
 DR
 PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 PT immune diseases including allograft rejection
 XX
 XX Example 7.1: Fig 18; 182pp: English.
 PS
 XX This polypeptide comprises the bispecific tetravalent antibody
 CC B17Ab1G10-B7-24H6. The molecule consists of 4 scfvs, i.e. 2
 CC anti B7.1 scfvs and 2 anti B7.2 scfvs (tetravalency). One single
 CC B17Ab is a homodimer of 2 identical molecules, each containing both
 CC an anti B7.1 and anti B7.2 scfv (bispecificity). An anti-B7.1 and
 CC and anti-B7.2 scfv are linked using a dimerisation domain (see
 CC W90219), which drives the homodimerisation of the molecule. DNA
 CC (see X01652) encoding the B17Ab has been constructed to allow
 CC expression of the B17Ab in transformed E. coli cells. The B17Ab
 CC cross-links, and/or cross-reacts, with the costimulatory molecules
 CC B7.1 and B7.2 that are expressed on the membrane of professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. The invention relates to such
 CC B7-binding molecules, methods for their production, and their use
 CC for treating or preventing diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).
 XX
 XX Sequence 556 AA:
 SQ

Query Match 67.7%; Score 63; DB 20; Length 556;
 Best Local Similarity 73.7%; Pred. NO. 0.068;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 LQLAESGVLVNPGXSDDL 20
 Db 312 VQIQESGGIVQPGSRRLL 330

RESULT 12
 W90217
 ID W90217 standard; Protein: 580 AA.
 AC
 XX W90217:
 XX
 XX 10-MAY-1999 (first entry)
 DT
 XX
 XX Bispecific tetravalent antibody B17AB7-24-IG10H6.
 DE
 XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
 XX CD86; T cell activation; inhibitor; graft versus host disease;
 KW transplant rejection; allograft rejection; autoimmune disease;
 KW allergy; therapy; human; bispecific tetravalent antibody; B17Ab;
 B17AB7-24-IG10H6.
 KM
 XX
 XX Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Chimeric - synthetic.
 XX

FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /note="peB signal peptide"
 FT Region 25..138
 FT /note="VH region anti B7.1 Mab"
 FT Peptide 139..153
 FT /note="(G4S3) flexible linker"
 FT Region 154..262
 FT /note="VL region anti B7.1 Mab"
 FT Region 263..273
 FT /note="human IgG3 hinge region"
 FT Domain 274..308
 FT /note="human IgG3 hinge region"
 FT Domain 309..319
 FT /note="helix-turn-helix dimerisation domain"
 FT Region 320..446
 FT /note="human IgG3 hinge domain"
 FT Peptide 447..461
 FT /note="VH region anti B7.2 Mab"
 FT Region 462..574
 FT /note="(G4S3)flexible linker"
 FT Peptide 575..580
 FT /note="VL region anti B7.2 Mab"
 FT /note="His6 tag"
 FT Misc-difference 261
 FT /note="encoded by CTC"
 FT Misc-difference 322..327
 FT /note="codons for these amino acids are not present in the DNA sequence for B17AB7-24-1G1-H6 provided in the specification"
 FT
 FT W09858965-A2.
 XX
 XX 30-DEC-1998.
 PD
 XX
 XX 22-JUN-1998: 98WO-EP03791.
 XX
 XX 20-JUN-1997: 97EP-0870092.
 PR
 XX (INNO-) INNOGENETICS NV.
 PA
 XX Bosman A, Buyse M, Lorre K, Sablon E;
 PI
 XX WPI: 1999-105615/09.
 DR
 XX N-PSDB: X01651.
 DR
 PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
 PT immune diseases including allograft rejection
 XX
 XX Example 7.1: Fig 16; 182pp: English.
 PS
 XX This polypeptide comprises the bispecific tetravalent antibody
 CC B17AB7-24-IG10H6. The molecule consists of 4 scfvs, i.e. 2
 CC anti B7.1 scfvs and 2 anti B7.2 scfvs (tetravalency). One single
 CC B17Ab is a homodimer of 2 identical molecules, each containing both
 CC an anti B7.1 and anti B7.2 scfv (bispecificity). An anti-B7.1 and
 CC and anti-B7.2 scfv are linked using a dimerisation domain (see
 CC W90219), which drives the homodimerisation of the molecule. DNA
 CC (see X01651) encoding the B17Ab has been constructed to allow
 CC expression of the B17Ab in transformed E. coli cells. The B17Ab
 CC cross-links, and/or cross-reacts, with the costimulatory molecules
 CC B7.1 and B7.2 that are expressed on the membrane of professional
 CC antigen-presenting cells, leading to the inhibition of antigen-
 CC specific T cell activation. The invention relates to such
 CC B7-binding molecules, methods for their production, and their use
 CC for treating or preventing diseases of the immune system, in
 CC particular graft rejection, graft versus host disease, allergy and
 CC autoimmune diseases (claimed).
 XX
 XX Sequence 580 AA:
 SQ

Query Match 67.7%; Score 63; DB 20; Length 580;

Best Local Similarity 73.7%; Pred. No. 0.072;
Matches 14; Conservative. 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGVLPQGXSDRL 20
:||| |||| ||||| |||
Db 24 vqlvesggdlvqpgsrrl 42

RESULT 13

W68169
ID W68169 standard; peptide; 20 AA.

XX W68169;

DT 06-OCR-1998 (first entry)

XX Human IgG heavy chain V-II region (HV3r) peptide.

XX Thrombolytic activity; soluble urokinase plasminogen activator receptor;

KW single chain urokinase type plasminogen activator; thromboembolism;

KW scuPA; IgG; regulator; fibrinolytic activity; fibrin clot;

KW myocardial infarction; cerebro-vascular event; pulmonary embolism;

XX deep vein thrombosis; immunoglobulin; human.

XX Homo sapiens.

XX MO9825641-A1.

XX 18-JUN-1998.

XX 09-DEC-1997; 97WO-1100402.

XX 09-DEC-1996; 96US-0032676.

XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

XX Hlgaz1 AA;

XX WPI: 1998-348262/30.

XX New compositions with thrombolytic activity for, e.g. treatment of

PT thromboembolism - comprise complex of single chain urokinase type

PT plasminogen activator, scuPA, and soluble urokinase plasminogen

PT activator receptor, scuPAR

XX Example 5; Page 25; 46pp; English.

XX This represents a peptide sequence from the human immunoglobulin G

CC (IgG) heavy chain, IG V-II region (HV3r). An IgG derived peptide

CC sequence having similarity to this sequence has a stimulating effect on

CC the fibrinolytic activity of the scuPA/suPAR complex of the invention.

CC The invention provides a thrombolytic therapeutic composition for the

CC treatment and prevention of a thromboembolic disorder associated with

CC the formation of fibrin clots. The composition comprises, as the active

CC ingredient, a complex of a single chain urokinase type plasminogen

CC activator (scuPA) and a soluble urokinase plasminogen activator receptor

CC (suPAR). The complex (scuPA/suPAR) has thrombolytic activity under

CC physiological conditions and in the presence of IgG, or of at least 1

CC IgG-derived peptide, and induces fibrinolysis of fibrin clots. The

CC compositions and complex are useful for the treatment or prevention of

CC thromboembolic disorders associated with the formation of fibrin clots,

CC especially myocardial infarctions, cerebro-vascular events, pulmonary

CC embolism and deep vein thrombosis.

XX Sequence 20 AA;

XX SQ

QY 2 LQLAESGGVLPQGXSDRL 20
:||| |||| ||||| |||

Best Local Similarity 73.7%; Score 62; DB 19; Length 20;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 2 vqlvesggdlvqpgsrrl 20

RESULT 14

W94487
ID W94487 standard; peptide; 20 AA.

XX W94487;

XX 21-APR-1999 (first entry)

XX Human adult heart zonulin N-terminal peptide.

XX Zonulin; mammalian tight junction; zonula occludens toxin; ZOT;

KW Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody;

KW intestinal mucosa; nasal mucosa; blood brain barrier.

XX Homo sapiens.

XX MO9852415-A1.

XX 26-NOV-1998.

XX 28-APR-1998; 98WO-US07636.

XX 21-MAY-1997; 97US-0859931.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Fasano A;

XX WPI: 1999-070123/06.

XX New purified zonulin - which is capable of reversibly opening

PT mammalian tight junctions, used for enhancing the delivery of agents

PT across intestinal and nasal mucosa and blood brain barrier

XX Claim 2; Page 45; 64pp; English.

XX The present invention describes pure zonulin which has an apparent

CC molecular weight of 47 kD, as determined by SDS-PAGE, which is

CC recognised by both anti-tau polyclonal antibody and by anti-zonula

CC occludens toxin (ZOT) polyclonal antibody, and is capable of reversibly

CC opening mammalian tight junctions. Zonulin proteins function as

CC physiological modulators of mammalian tight junctions. They can be used

CC for enhancing the absorption of therapeutic agents across tight

CC junctions of intestinal and nasal mucosa and across tight junctions of

CC e.g. lidocaine, adenosine, dobutamine, dopamine, epinephrine,

CC norepinephrine, phenolamine, doxapram, alfentanil, dezocin, nalbuphine,

CC buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurine,

CC vincristine, vincristine, methicillin, mezlocillin, piperacillin,

CC cefotaxime, ceftriaxone, cefmetazole and aztreonam, a hormone e.g.

CC testosterone, nandrolone, menotropins, insulin, urofollitropin,

CC interferon-alpha, interferon-beta, interferon-gamma, interleukin-1

CC (IL-1), IL-2, IL-4, IL-8, polyvalent IgG, specific IgG, IgM, or IgM.

CC The proteins can also be used for the production of antibodies which can

CC be used to assay for zonulin in body tissue or fluids, or in affinity-

CC purification of zonulin. The present sequence represents an N-terminal

CC peptide of zonulin.

XX Sequence 20 AA;

XX SQ

QY 2 LQLAESGGVLPQGXSDRL 20
:||| |||| ||||| |||

Best Local Similarity 73.7%; Score 62; DB 20; Length 20;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 2 vqlvesggdlvqpgsrrl 20

RESULT 15

Y79130
ID Y79130 standard; Peptide; 20 AA.

XX
AC Y79130;

XX
DE 05-JUN-2000 (first entry)

XX
DE Human adult heart zonulin N-terminal sequence.

XX
KW Zonulin; antagonist; zonula occludens toxin receptor;

KW human; blood-brain barrier; antiinflammatory;

KW gastrointestinal inflammation; therapy.

XX
OS Homo sapiens.

XX
PN WC020007609-A1.

XX
PD 17-FEB-2000.

XX
PE 28-JUL-1999; 99WO-US16683.

XX
PR 03-AUG-1998; 98US-0127815.

XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.

XX
PI Faasano A;

XX
DR WPI: 2000-205565/18.

XX
PT New peptide antagonist of zonulin useful as antiinflammatory agent for

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,

PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis

XX
PS Example 3; Fig 6; 69pp; English.

XX
CC The present sequence is that of the N-terminal region of adult

CC human heart zonulin. The N-terminal sequences of human adult and

CC foetal zonulins (see Y79130-36) were compared with Vibrio cholerae

CC zonula occludens toxin (ZOT) to identify a common motif thought

CC to be involved in receptor binding. Peptide antagonists (see

CC Y79105-29) based on this motif are useful as antiinflammatory

CC agents for treatment of gastrointestinal inflammation, and for

CC treatment of conditions associated with breakdown of the blood-brain

CC barrier.

XX
SQ Sequence 20 AA:

XX
QY Query Match 66.7%; Score 62; DB 21; Length 20;

XX
Best Local Similarity 73.7%; Pred. NO. 0.0027;

XX
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

XX
Db 2 LQLAESGGVIVQPGXSRL 20

XX
2 vqlvesgggllvqpggslrl 20

XX
Search completed: June 13, 2001, 14:25:46

XX
Job time: 665 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:11 ; Search time 87.97 Seconds

(without alignments)
15,624 Million cell updates/sec

Title: PCT-US01-05825A-27

Perfect score: 93

Sequence: 1 MLOAESGVLVQPGXSRL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	73.1	116	1	IGHUZP
2	66	71.0	98	2	S26929
3	66	71.0	120	2	S41111
4	64	68.8	147	2	PH0120
5	64	68.8	148	2	PH0116
6	64	68.8	148	2	PH0115
7	64	68.8	148	2	PH0118
8	64	68.8	148	2	PH0121
9	64	68.8	148	2	PH0119
10	64	68.8	148	2	PH0117
11	62	66.7	97	2	PH0872
12	62	66.7	97	2	S26890
13	62	66.7	97	2	S46462
14	62	66.7	98	2	PL0121
15	62	66.7	98	2	PL0123
16	62	66.7	98	2	S26896
17	62	66.7	98	2	S29545
18	62	66.7	98	2	S26932
19	62	66.7	98	2	S26932
20	62	66.7	98	2	S26891
21	62	66.7	98	2	S26894
22	62	66.7	98	2	S26933
23	62	66.7	98	2	S26934
24	62	66.7	100	2	S69896
25	62	66.7	100	2	S26925
26	62	66.7	100	2	S26926
27	62	66.7	114	2	S36280
28	62	66.7	116	1	M3HUGL
29	62	66.7	117	2	S17079

ALIGNMENTS

30	62	66.7	117	2	A34964	Ig heavy chain pre
31	62	66.7	117	2	S36259	Ig heavy chain V r
32	62	66.7	117	2	S21980	Ig heavy chain V-g
33	62	66.7	118	2	S31121	Ig heavy chain - h
34	62	66.7	119	1	G1HUTE	Ig heavy chain V-I
35	62	66.7	119	2	S31107	Ig heavy chain - h
36	62	66.7	120	1	M3HDBW	Ig heavy chain V-I
37	62	66.7	120	1	G1HDB	Ig heavy chain V-I
38	62	66.7	121	2	S26798	Ig heavy chain V r
39	62	66.7	121	2	S31118	Ig heavy chain - h
40	62	66.7	121	2	S31104	Ig heavy chain (su
41	62	66.7	123	2	S31509	Ig heavy chain - h
42	62	66.7	123	2	S26794	Ig heavy chain V r
43	62	66.7	123	2	S30532	Ig heavy chain V r
44	62	66.7	124	2	S20775	Ig heavy chain V r
45	62	66.7	125	2	S30531	Ig heavy chain V r

RESULT 1
IGHUZP
Ig heavy chain V-IIT region (Zap) - human (tentative sequence)
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000
C:Accession: A02061
R:Capra, J.D.; Kehoe, J.M.
Proc. Natl. Acad. Sci. U.S.A. 71, 845-848, 1974
A:Title: Variable region sequences of five human immunoglobulin heavy chains of the V
A:Reference number: A93794; M0ID:74142702
A:Accession: A02061
A:Molecule type: protein
A:Residues: 1-116 <CAP>
C:Comment: This chain was isolated from an Iga1 myeloma protein.
C:Genetics:
A:Gene: GDB:IGHV4
A:Cross-references: GDB:128528; OMIM:147070
A:Map position: 14q32.33-14q32.33
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:22-96/Disulfide bonds: #status predicted

Query Match 73.1%; Score 68; DB 1; Length 116;
Best Local Similarity 73.7%; Pred. No. 0.00036;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 L0LAESGVLVQPGXSRL 20
DB 2 VL0VESGALVQPGSGRL 20

RESULT 2
S26929
Ig heavy chain V region (DP-33) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
R:Tromlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; M0ID:93021117
A:Accession: S26929
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12335; NID:g32889; PIDN:CAA78205.1; PID:g32890
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 71.0%; Score 66; DB 2; Length 98;
Best Local Similarity 73.7%; Pred. No. 0.00064;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGXSRL 20
DB 2 VOLVESGGVVPQGSRL 20

RESULT 3
S44111
Ig heavy chain V-D-J region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S44111
R:Zelenetz, R.E.; Zhu, D.; Ovecke, M.; Winter, G.; Hamblin, T.J.; Stevenson, F.K.
submitted to the EMBL Data Library, March 1994
A:Description: Idiotype vaccination against human B-cell lymphoma: rescue of variable
A:Reference number: S44105
A:Accession: S44105
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <HAW>
A:Cross-references: EMBL:Z31387; NID:9472965; PIDN:CA83262.1; PID:9940522
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:13-98/Domain: Immunoglobulin homology <IMM>

Query Match 71.0%; Score 66; DB 2; Length 120;
Best Local Similarity 73.7%; Pred. No. 0.00079;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGXSRL 20
DB 2 VOLVESGGVVPQGSRL 20

RESULT 4
PH0120
Ig heavy chain precursor V-D-J region (JP-FL-6) - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-1996
C:Accession: PH0120
R:Zelenetz, A.D.; Chen, T.T.; Levy, R.
J. Exp. Med. 173, 197-207, 1991
A:Title: Histologic transformation of follicular lymphoma to diffuse lymphoma represents
A:Reference number: PH0115; MUID:91086845
A:Accession: PH0120
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-147 <ZEL>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-147/Product: Ig heavy chain #status predicted <MAT>
F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 68.8%; Score 64; DB 2; Length 147;
Best Local Similarity 73.7%; Pred. No. 0.0021;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGXSRL 20
DB 21 LQLVESGGVLPQGSRL 39

RESULT 5
PH0116
Ig heavy chain precursor V-D-J region (JP-FL-2) - human
C:Species: Homo sapiens (man)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-1996
C:Accession: PH0116
R:Zelenetz, A.D.; Chen, T.T.; Levy, R.
J. Exp. Med. 173, 197-207, 1991
A:Title: Histologic transformation of follicular lymphoma to diffuse lymphoma represe
A:Reference number: PH0115; MUID:91086845
A:Accession: PH0116
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-148 <ZEL>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-148/Product: Ig heavy chain #status predicted <MAT>
F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 68.8%; Score 64; DB 2; Length 148;
Best Local Similarity 73.7%; Pred. No. 0.0021;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGXSRL 20
DB 21 LQLVESGGVLPQGSRL 39

RESULT 6
PH0118
Ig heavy chain precursor V-D-J region (JP-FL-4) - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-1996
C:Accession: PH0118
R:Zelenetz, A.D.; Chen, T.T.; Levy, R.
J. Exp. Med. 173, 197-207, 1991
A:Title: Histologic transformation of follicular lymphoma to diffuse lymphoma represe
A:Reference number: PH0115; MUID:91086845
A:Accession: PH0118
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-148 <ZEL>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-148/Product: Ig heavy chain #status predicted <MAT>
F:34-117/Domain: Immunoglobulin homology <IMM>

Query Match 68.8%; Score 64; DB 2; Length 148;
Best Local Similarity 73.7%; Pred. No. 0.0021;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGXSRL 20
DB 21 LQLVESGGVLPQGSRL 39

RESULT 7
PH0115
Ig heavy chain precursor V-D-J region (JP-FL-1) - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-1996
C:Accession: PH0115
R:Zelenetz, A.D.; Chen, T.T.; Levy, R.
J. Exp. Med. 173, 197-207, 1991
A:Title: Histologic transformation of follicular lymphoma to diffuse lymphoma represe
A:Reference number: PH0115; MUID:91086845
A:Accession: PH0115
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-148 <ZEL>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-148/Product: Ig heavy chain #status predicted <MAT>
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 68.8%; Score 64; DB 2; Length 148;
Best Local Similarity 73.7%; Pred. No. 0.0021;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 21 LQLVESGGVLVOPGGSRL 39

RESULT 8

Ig heavy chain precursor V-D-J region (JP-FL-7) - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-1996
C:Accession: PH0121

R:Zelenetz, A.D.; Chen, T.T.; Levy, R.

J. Exp. Med. 173, 197-207, 1991
A:Title: Histologic transformation of follicular lymphoma to diffuse lymphoma represents
A:Reference number: PH0115; MUID:91086845
A:Accession: PH0121

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-148 <ZEL>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-148/Product: Ig heavy chain #status predicted <MAT>

F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 68.8%; Score 64; DB 2; Length 148;
Best Local Similarity 73.7%; Pred. No. 0.0021;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 21 LQLVESGGVLVOPGGSRL 39

RESULT 9

Ig heavy chain precursor V-D-J region (JP-FL-5) - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-1996
C:Accession: PH0119

R:Zelenetz, A.D.; Chen, T.T.; Levy, R.

J. Exp. Med. 173, 197-207, 1991
A:Title: Histologic transformation of follicular lymphoma to diffuse lymphoma represents
A:Reference number: PH0115; MUID:91086845
A:Accession: PH0119

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-148 <ZEL>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-148/Product: Ig heavy chain #status predicted <MAT>

F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 68.8%; Score 64; DB 2; Length 148;
Best Local Similarity 73.7%; Pred. No. 0.0021;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 21 LQLVESGGVLVOPGGSRL 39

RESULT 10
PH0117
Ig heavy chain precursor V-D-J region (JP-FL-3) - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Aug-1996
C:Accession: PH0117

R:Zelenetz, A.D.; Chen, T.T.; Levy, R.
J. Exp. Med. 173, 197-207, 1991
A:Title: Histologic transformation of follicular lymphoma to diffuse lymphoma represe
A:Reference number: PH0115; MUID:91086845
A:Accession: PH0117

A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-148 <ZEL>

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-148/Product: Ig heavy chain #status predicted <MAT>

F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 68.8%; Score 64; DB 2; Length 148;
Best Local Similarity 73.7%; Pred. No. 0.0021;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 21 LQLVESGGVLVOPGGSRL 39

RESULT 11

Ig heavy chain V region (anti-DNA, IIT-3R) - human (fragment)
PH0872
C:Species: Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 16-Aug-1996
C:Accession: PH0872

R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiot
A:Reference number: PH0862; MUID:92078875
A:Accession: PH0872

A:Molecule type: DNA

A:Residues: 1-97 <MAN>

A:Note: a "G" was inserted at the position of between 72nd and 73rd of the DNA sequen
C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that be
C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:14-97/Domain: immunoglobulin homology <IMM>

F:30-35/Region: complementarity-determining 1

F:49-66/Region: complementarity-determining 2

Query Match 66.7%; Score 62; DB 2; Length 97;
Best Local Similarity 73.7%; Pred. No. 0.0029;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 1 VQLVESGGVLVOPGGSRL 19

RESULT 12

Ig heavy chain V region (DP-48) - human (fragment)
S26890
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26890

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o

A:Reference number: S26885; MUID:93021117
A:Accession: S26890

A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-97 <TOM>
 A:Cross-references: EMBL:Z12348; NID:932916; PIDN:CAA78218.1; PID:932917
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:15-97/Domain: Immunoglobulin homology <IMM>

Query Match 66.7%; Score 62; DB 2; Length 97;
 Best Local Similarity 73.7%; Pred. No. 0.0029;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGVLVQPGXSRL 20
 :|| |||| |||| | ||
 Db 2 VQLVESGGGLVQPGGSRL 20

RESULT 13
 S46462
 Ig heavy chain V region (VAC-5) - human
 C:Species: Homo sapiens (man)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
 R:Accession: S46462
 R:Cook, G.P.; Tomlinson, I.M.; Walter, G.; Riethman, H.; Carter, N.P.; Buluwela, L.; Wit
 Nature Genet. 7, 162-168, 1994
 A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the telom
 A:Reference number: S46460; MUID:95004581
 A:Accession: S46462
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-97 <COO>
 A:Cross-references: EMBL:Z27504; NID:9505430; PIDN:CAA81824.1; PID:9505431
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: heterotetramer; Immunoglobulin
 F:15-97/Domain: Immunoglobulin homology <IMM>

Query Match 66.7%; Score 62; DB 2; Length 97;
 Best Local Similarity 73.7%; Pred. No. 0.0029;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGVLVQPGXSRL 20
 :|| |||| |||| | ||
 Db 2 VQLVESGGGLVQPGGSRL 20

RESULT 14
 PL0121
 Ig heavy chain V-III region (TD-Vp) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Aug-1996
 C:Accession: PL0121
 R:Bird, J.; Gallili, N.; Link, M.; Stiles, D.; Sklar, J.
 J. Exp. Med. 168, 229-245, 1988
 A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin
 A:Reference number: PL0116; MUID:88286083
 A:Accession: PL0121
 A:Molecule type: mRNA
 A:Residues: 1-98 <BIR>
 A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
 A>Note: the sequence shows the V region (TD-Vp) from one of five DNA rearrangements from
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: acute lymphoblastic leukemia; heterotetramer; Immunoglobulin
 F:15-98/Domain: Immunoglobulin homology <IMM>
 F:31-35/Region: complementarity-determining 1
 F:49-65/Region: complementarity-determining 2

Query Match 66.7%; Score 62; DB 2; Length 98;
 Best Local Similarity 73.7%; Pred. No. 0.003;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGVLVQPGXSRL 20

Db 2 VQLVESGGGLVQPGGSRL 20
 :|| |||| |||| | ||

RESULT 15
 PL0123
 Ig heavy chain V-III region (TD-Vr) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
 R:Accession: PL0123; S26897
 R:Bird, J.; Gallili, N.; Link, M.; Stiles, D.; Sklar, J.
 J. Exp. Med. 168, 229-245, 1988
 A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobu
 A:Reference number: PL0116; MUID:88286083
 A:Accession: PL0123
 A:Molecule type: mRNA
 A:Residues: 1-98 <BIR>
 A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
 A>Note: the sequence shows the V region (TD-Vr) from a nonproductive DNA rearrangemen
 R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
 J. Mol. Biol. 227, 776-798, 1992
 A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
 A:Reference number: S26885; MUID:93021117
 A:Accession: S26897
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-98 <TOM>
 A:Cross-references: EMBL:Z12354; NID:932930; PIDN:CAA78224.1; PID:932931
 C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
 C:Keywords: acute lymphoblastic leukemia; heterotetramer; Immunoglobulin
 F:15-98/Domain: Immunoglobulin homology <IMM>
 F:31-35/Region: complementarity-determining 1
 F:49-65/Region: complementarity-determining 2

Query Match 66.7%; Score 62; DB 2; Length 98;
 Best Local Similarity 73.7%; Pred. No. 0.003;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGVLVQPGXSRL 20
 :|| |||| |||| | ||
 Db 2 VQLVESGGGLVQPGGSRL 20

Search completed: June 13, 2001, 14:23:12
 Job time: 743 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM proteain - proteain search, using sw model

Run on: June 13, 2001, 14:30:37 ; Search time 51.57 Seconds
(without alignments)
13.285 Million cell updates/sec

Title: PCT-US01-05825A-27
Perfect score: 93
Sequence: 1 MLOLAESGCVLPQPGXSDRL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	73.1	116	1 HV30_HUMAN	P01778 homo sapien
2	62	66.7	116	1 HV31_HUMAN	P01781 homo sapien
3	62	66.7	119	1 HV32_HUMAN	P01777 homo sapien
4	62	66.7	120	1 HV33_HUMAN	P01766 homo sapien
5	62	66.7	120	1 HV34_HUMAN	P01782 homo sapien
6	61	65.6	115	1 HV3D_HUMAN	P01765 homo sapien
7	61	65.6	116	1 HV05_CARAU	P19181 carassius a
8	61	65.6	116	1 HV3R_HUMAN	P01779 homo sapien
9	61	65.6	117	1 HV3C_HUMAN	P01764 homo sapien
10	61	65.6	117	1 HV3O_HUMAN	P01776 homo sapien
11	61	65.6	119	1 HV3K_HUMAN	P01774 homo sapien
12	61	65.6	119	1 HV3N_HUMAN	P01775 homo sapien
13	59	63.4	118	1 HV16_MOUSE	P01783 mus musculu
14	59	63.4	118	1 HV3V_HUMAN	P80419 homo sapien
15	59	63.4	126	1 HV3K_HUMAN	P01772 homo sapien
16	58	62.4	114	1 HV01_CANFA	P01784 canis faml
17	58	62.4	115	1 HV3E_HUMAN	P01767 homo sapien
18	58	62.4	115	1 HV3S_HUMAN	P01780 homo sapien
19	58	62.4	117	1 HV17_MOUSE	P01786 mus musculu
20	58	62.4	122	1 HV20_MOUSE	P01789 mus musculu
21	58	62.4	122	1 HV21_MOUSE	P01790 mus musculu
22	58	62.4	123	1 HV18_MOUSE	P01787 mus musculu
23	58	62.4	123	1 HV19_MOUSE	P01788 mus musculu
24	58	62.4	123	1 HV22_MOUSE	P01792 mus musculu
25	58	62.4	123	1 HV24_MOUSE	P01793 mus musculu
26	58	62.4	123	1 HV25_MOUSE	P01794 mus musculu
27	58	62.4	123	1 HV26_MOUSE	P01763 homo sapien
28	56	60.2	114	1 HV3S_HUMAN	P01770 homo sapien
29	56	60.2	119	1 HV3I_HUMAN	P18528 mus musculu
30	55	59.1	98	1 HV37_MOUSE	P18528 mus musculu
31	55	59.1	113	1 HV27_MOUSE	P01797 mus musculu
32	55	59.1	113	1 HV28_MOUSE	P01798 mus musculu
33	55	59.1	113	1 HV29_MOUSE	P01798 mus musculu

34	55	59.1	113	1 HV30_MOUSE	P01799 mus musculu
35	55	59.1	113	1 HV31_MOUSE	P01800 mus musculu
36	55	59.1	113	1 HV34_MOUSE	P01803 mus musculu
37	55	59.1	115	1 HV32_MOUSE	P01801 mus musculu
38	55	59.1	115	1 HV33_MOUSE	P18525 mus musculu
39	55	59.1	117	1 HV54_MOUSE	P18526 mus musculu
40	55	59.1	117	1 HV55_MOUSE	P18526 mus musculu
41	55	59.1	122	1 HV3A_HUMAN	P01762 homo sapien
42	55	59.1	142	1 HV01_RAT	P01805 rattus norv
43	55	59.1	144	1 HV36_MOUSE	P01795 mus musculu
44	54	58.1	116	1 HV36_MOUSE	P01806 mus musculu
45	54	58.1	117	1 HV02_CANFA	P01785 canis faml

ALIGNMENTS

```

RESULT 1
HV30_HUMAN          STANDARD;          PRT; 116 AA.
ID HV30_HUMAN
AC P01778;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION ZAP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kenoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains
RT of the VH3 subgroup: definitive identification of four heavy chain
RT hypervariable regions."
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGAL MYELOMA
CC PROTEIN.
CC PIR; A02061; AIH02P.
DR HSSP; P01772; 2IG2.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; 1g; 1.
KW Immunoglobulin V region.
FT NON_TER
SQ
SEQUENCE 116 AA; 12582 MW; 892F8C217CEC9865 CRC64;

Query Match 73.1%; Score 68; DB 1; Length 116;
Best Local Similarity 73.7%; Pred. NO. 0.0001;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LQLAESGCVLPQPGXSDRL 20
Db 2 VQLVESGALVLPQPGSGRL 20

RESULT 2
HV31_HUMAN          STANDARD;          PRT; 116 AA.
ID HV31_HUMAN
AC P01781;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION GAL.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=75059123; PubMed=4803843;
RA Metanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;

```

"The primary structure of a monoclonal IgM-immunoglobulin (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-type), subgroup H III. Architecture of the complete IgM-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
RN [2]
RP REVISION TO THE COMPOSITION OF 28-33.
RA Hilschmann N.;
RU Submitted (JUN-1975) to the PIR data bank.
CC -I- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S MACROGLOBULIN.
DR PIR: A02064; M3HUGL.
DR HSSP: P01772; 2IG2.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9AAAAA1282 CRC64;

Query Match 66.7%; Score 62; DB 1; Length 116;
Best Local Similarity 73.7%; Pred. No. 0.0011;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQAESGGVLVOPGXSRL 20
DB 2 VOLVESGGDLVOPGSRRL 20

RESULT 3
HV3U_HUMAN STANDARD; PRT; 119 AA.
AC P01777;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kehoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains of the VH3 subgroup: definitive identification of four heavy chain hypervariable regions.";
RT Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA PROTEIN.
CC PIR: A02060; GIHUTE.
DR HSSP: P01772; 2IG2.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 119 119
SQ SEQUENCE 119 AA; 12802 MW; 7E24DC852C7290A9 CRC64;

Query Match 66.7%; Score 62; DB 1; Length 119;
Best Local Similarity 73.7%; Pred. No. 0.0011;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQAESGGVLVOPGXSRL 20
DB 2 VOLVESGGDLVOPGSRRL 20

RESULT 4
HV3E_HUMAN STANDARD; PRT; 120 AA.
AC P01766;
DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION BRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77117674; PubMed=65324;
RA Capra J.D., Hopper J.E.;
RT "Comparative studies on monotypic IgM lambda and IgG kappa from an individual patient. III. The complete amino acid sequence of the VH region of the IgM paraprotein.";
RT Immunohemistry 13:995-999(1976).
CC -I- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM TYPE.
CC PIR: A02049; M3HUBW.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13327 MW; D3F0428F7C2E6410 CRC64;

Query Match 66.7%; Score 62; DB 1; Length 120;
Best Local Similarity 73.7%; Pred. No. 0.0011;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQAESGGVLVOPGXSRL 20
DB 2 VOLVESGGDLVOPGSRRL 20

RESULT 5
HV3U_HUMAN STANDARD; PRT; 120 AA.
AC P01782;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION DOB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=80020921; PubMed=114209;
RA Steiner L.A., Garcia Pardo A., Margolies M.N.;
RT "Amino acid sequence of the heavy-chain variable region of the crystallizable human myeloma protein Dob.";
RT Biochemistry 18:4068-4080(1979).
RN [2]
RP CRYSTALLIZATION.
RX MEDLINE=80020920; PubMed=114208;
RA Steiner L.A., Lopes A.D.;
RT "The crystallizable human myeloma protein Dob has a hinge-region deletion.";
RT Biochemistry 18:4054-4067(1979).
CC -I- MISCELLANEOUS: THIS GAMMA-1 MYELOMA PROTEIN HAS A DELETION IN THE HINGE REGION. THERE ARE NO LIGHT-HEAVY OR INTER-HEAVY CHAIN DISULFIDE BONDS.
CC PIR: A02065; GIHUBD.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;

Query Match 66.7%; Score 62; DB 1; Length 120;

Best Local Similarity 73.7%; Pred. No. 0.0011;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAESGGLVOPGXSRL 20
DB 2 VOLLESGGLVOPGXSRL 20

RESULT 6

HV3D_HUMAN STANDARD; PRT; 115 AA.
AC P01765;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.

RX MEDLINE-78005528; PubMed-409716;
RA Wang A.-C., Wang I.Y., Fudenberg H.H.;
RT "Immunoglobulin structure and genetics. Identity between variable regions of a mu and a gamma2 chain."
RL J. Biol. Chem. 252:7192-7199(1977).
CC -1- MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS OF IGM AND IG22 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO IDENTICAL.
CC PIR: A02048; H3HUTL.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 115
SQ SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 CRC64;

Query Match

Best Local Similarity 73.7%; Score 61; DB 1; Length 115;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAESGGLVOPGXSRL 20
DB 2 VOLLESGGLVOPGXSRL 20

RESULT 7

HV05_CARAU STANDARD; PRT; 116 AA.
AC P19181;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 5A PRECURSOR.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
OX NCBI_Taxid=7957;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE-88144476; PubMed-3125551;
RA Wilson M.R., Middleton D., Watt G.W.;
RT "Immunoglobulin heavy chain variable region gene evolution: structure and family relationships of two genes and a pseudogene in a teleost fish."
RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region; Signal.

FT SIGNAL 1 19
FT CHAIN 20 116 IG HEAVY CHAIN V REGION 5A.
FT DOMAIN 20 49 FRAMEWORK 1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 55 68 FRAMEWORK 2.
FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 85 116 FRAMEWORK 3.
FT DISULF 41 114 BY SIMILARITY.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;

Query Match

Best Local Similarity 68.4%; Score 61; DB 1; Length 116;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAESGGLVOPGXSRL 20
DB 21 VOLLESGGLVOPGXSRL 39

RESULT 8

HV3R_HUMAN STANDARD; PRT; 116 AA.
AC P01779;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.

RX MEDLINE-74142702; PubMed-4522793;
RA Capra J.D., Kehoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains of the VH3 subgroup: definitive identification of four heavy chain RT hypervariable regions."
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGAL MYELOMA PROTEIN.
CC PIR: A02062; A1HUTV.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12431 MW; EB705F53A963F0C CRC64;

Query Match

Best Local Similarity 73.7%; Score 61; DB 1; Length 116;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAESGGLVOPGXSRL 20
DB 2 VOLLESGGLVOPGXSRL 20

RESULT 9

HV3C_HUMAN STANDARD; PRT; 117 AA.
AC P01764;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION VH26 PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-81101090; PubMed-6450418;
 RA Mathysens G., Rabbits T.H.;
 RT "Structure and multiplicity of genes for the human immunoglobulin
 heavy chain variable region";
 RT Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J00236; AAA53516.1; -
 DR EMBL: M35415; AAA58735.1; -
 DR PIR: A02047; H3H026;
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1;
 DR Immunoglobulin V region; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 117 IG HEAVY CHAIN V-III REGION VH26.
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA: 12582 MW: 826733F1A3CB0F1 CRC64;

Query Match 65.6%; Score 61; DB 1; Length 117;
 Best Local Similarity 73.7%; Pred. No. 0.0016;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGGVLPQGSDDL 20
 DB 21 VQLESGGVLVPGGSLRL 39

RESULT 10
 HV30_HUMAN STANDARD; PRT; 117 AA.
 AC P01776;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION WAS.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=74142702; PubMed=4522793;
 RA Capra J.D., Kehoe J.M.;
 RT "Variable region sequences of five human immunoglobulin heavy chains
 of the Vh3 subgroup: definitive identification of four heavy chain
 hypervariable regions";
 RT Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
 PROTEIN.
 DR PIR: A02059; GIHWMS.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1;
 DR Immunoglobulin V region.
 FT NON_TER 117 117
 SQ SEQUENCE 117 AA: 13091 MW: 201DEF0E1E53D9BF CRC64;

Query Match 65.6%; Score 61; DB 1; Length 117;
 Best Local Similarity 73.7%; Pred. No. 0.0016;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 OY 2 LQLAESGGVLPQGSDDL 20
 DB 21 VQLESGGVLVPGGSLRL 39

DB 2 VQLESGGVLVPGGSLRL 20
 RESULT 11
 HV3M_HUMAN STANDARD; PRT; 119 AA.
 AC P01774;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION POM.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75046755; PubMed=4139708;
 RA Capra J.D., Kehoe J.M.;
 RT "Structure of antibodies with shared idiotypic: the complete sequence
 of the heavy chain variable regions of two immunoglobulin M
 anti-gamma globulins";
 RT Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
 GLOBULIN ACTIVITY.
 DR PIR: A02057; M3H0M.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1;
 DR Immunoglobulin V region.
 FT VARIANT 54 54
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA: 12953 MW: 2E018AF4DCEB2610 CRC64;

Query Match 65.6%; Score 61; DB 1; Length 119;
 Best Local Similarity 73.7%; Pred. No. 0.0016;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGGVLPQGSDDL 20
 DB 2 VQLESGGVLVPGGSLRL 20

RESULT 12
 HV3N_HUMAN STANDARD; PRT; 119 AA.
 AC P01775;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION LAY.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75046755; PubMed=4139708;
 RA Capra J.D., Kehoe J.M.;
 RT "Structure of antibodies with shared idiotypic: the complete sequence
 of the heavy chain variable regions of two immunoglobulin M
 anti-gamma globulins";
 RT Proc. Natl. Acad. Sci. U.S.A. 71:4032-4036(1974).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM IGM WITH ANTI-GAMMA
 GLOBULIN ACTIVITY.
 DR PIR: A02058; M3H0L.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 19; 1;
 DR Immunoglobulin V region.

FT NON_TER 119 119
SQ SEQUENCE 119 AA; 12858 MW; D6338098794DCF5E CRC64;

Query Match 65.6%; Score 61; DB 1; Length 119;
Best Local Similarity 73.7%; Pred. No. 0.0016;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGVLVQPGXSRL 20
DB 2 VOLVESGGVLVQPGGSLRL 20

RESULT 13
HV16_MOUSE
ID HV16_MOUSE STANDARD; PRT; 136 AA.

AC P01783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION MOPC 21 PRECURSOR (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
[1]
RN RP
RX MEDLINE=61234546; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K., Baltimore D.;

RT "Heavy chain variable region contribution to the Npb family of antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:635-637(1981).
[2]

RN RP
RX MEDLINE=77100368; PubMed=401950;
RA Adetunbo K., Milstein C., Secher D.S.;
RT "Molecular analysis of spontaneous somatic mutants.";
RL Nature 265:299-304(1977).
[1]

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL: J00522; AAD15290.1; -;
DR PIR: A02066; GIMS21;
DR InterPro: IPR003006; -;
DR Pfam: PF00047; 1g; 1;
KW Immunoglobulin V region; signal.
FT SIGNAL 1 16
FT CHAIN 17 136
FT DOMAIN 115 119
FT DOMAIN 120 136
FT DISULFID 38 112
FT CONFLICT 75 78
FT CONFLICT 89 90
FT CONFLICT 115 115
FT CONFLICT 120 120
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15071 MW; 2276A98DBDEF7016 CRC64;

Query Match 65.6%; Score 61; DB 1; Length 136;
Best Local Similarity 68.4%; Pred. No. 0.0019;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 LQLAESGVLVQPGXSRL 20
DB 2 VOLVESGGVLVQPGGSLRL 20

DB 18 VOLVESGGVLVQPGGSRKL 36

RESULT 14
HV3V_HUMAN
ID HV3V_HUMAN STANDARD; PRT; 118 AA.
AC P80419;

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION GAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
[1]

RN RP
RX MEDLINE=95255298; PubMed=7737190;
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferrari G.;
RT "Characterization of the two unique human anti-flavin monoclonal antibodies.";
RL Eut. J. Biochem. 228: 886-893(1995).
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR003006; -;
DR Pfam: PF00047; 1g; 1;
KW Immunoglobulin V region.
FT NON_TER 118 118
SQ SEQUENCE 118 AA; 13087 MW; 6C21D810ED1B6D1F CRC64;

Query Match 63.4%; Score 59; DB 1; Length 118;
Best Local Similarity 68.4%; Pred. No. 0.0035;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGVLVQPGXSRL 20
DB 2 VOLVESGGVLVQPGGSLRL 20

RESULT 15
HV3K_HUMAN
ID HV3K_HUMAN STANDARD; PRT; 126 AA.
AC P01772;

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION KOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
[1]

RN RP
RX MEDLINE=83289131; PubMed=688494;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
[2]

RN RP
RX X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
MEDLINE=81072295; PubMed=7441755;
RA Marquart M., Deisenhofer J., Huber R., Palm W.;
RT "Crystallographic refinement and atomic models of the intact immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A and 1.0-A resolution.";
RL J. Mol. Biol. 141:369-391(1980).
PIR: A02055; G1HUKL.
PDB: 2FBJ; 12-JUL-89.
PDB: 2IG2; 12-JUL-89.
DR InterPro: IPR003006; -;
DR Pfam: PF00047; 1g; 1;
KW Immunoglobulin V region; 3D-structure.
MOD_RES 1 1
PYRROLIDONE CARBOXYLIC ACID.

Query Match 63.4%; Score 59; DB 1; Length 118;
Best Local Similarity 68.4%; Pred. No. 0.0035;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 LQLAESGVLVQPGXSRL 20
DB 2 VOLVESGGVLVQPGGSLRL 20

```
FT DISULFID 22 96
FT DISULFID 105 110
FT STRAND 3 7
FT STRAND 11 12
FT TURN 14 13
FT STRAND 18 25
FT HELIX 29 31
FT STRAND 34 39
FT TURN 41 42
FT STRAND 46 51
FT TURN 53 54
FT STRAND 58 60
FT TURN 62 67
FT STRAND 68 73
FT TURN 74 77
FT STRAND 78 83
FT HELIX 88 90
FT STRAND 92 99
FT STRAND 106 106
FT TURN 107 108
FT STRAND 109 109
FT STRAND 113 113
FT STRAND 120 124
FT NON_TER. 126 126
SQ SEQUENCE 126 AA; 13718 MM; EAD71B52B16F8776 CRC64;
```

```
Query Match 63.4%; Score 59; DB 1; Length 126;
Best Local Similarity 68.4%; Pred. No. 0.0038;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 2 LQAESGGVLPQPGXSRL 20
Db 2 VQVESGGGVQPGRSRL 20
```

Search completed: June 13, 2001, 14:30:37
Job time: 527 sec

DT 01-MAY-2000 (TREMBLERL. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLERL. 15, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus".
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035042; AAD56278.1; -.
 DR HSSP: P01772; 2F84.
 DR INTERPRO: IPR003006; -.
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944BD5CCA CRC64;

Query Match 66.7%; Score 62; DB 4; Length 118;
 Best Local Similarity 73.7%; Pred. No. 0.012;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAEAGGVVOPGXSDRL 20
 DB 2 VOLVESGGVOPGSLRL 20
 :||| ||||| ||||| |||

RESULT 3
 ID Q9ULB6 PRELIMINARY; PRT; 95 AA.
 AC Q9ULB6;
 DT 01-MAY-2000 (TREMBLERL. 13, Created)
 DT 01-MAY-2000 (TREMBLERL. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLERL. 15, Last annotation update)
 DE IMMUNOGLOBULIN HEAVY CHAIN (FRAGMENT).
 GN VH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tange Y., Kayano H.;
 RT "Human VH gene sequence".
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB035268; BAA87067.1; -.
 DR HSSP: P01772; 2F84.
 DR INTERPRO: IPR003006; -.
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 95 AA; 10527 MW; 90A8C6D16D22574A CRC64;

Query Match 65.6%; Score 61; DB 4; Length 95;
 Best Local Similarity 73.7%; Pred. No. 0.013;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAEAGGVVOPGXSDRL 20
 DB 1 VOLVESGGVOPGSLRL 19
 :||| ||||| ||||| |||

RESULT 4
 ID Q9UL90 PRELIMINARY; PRT; 113 AA.
 AC Q9UL90;
 Q9UL90;

DT 01-MAY-2000 (TREMBLERL. 13, Created)
 DT 01-MAY-2000 (TREMBLERL. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLERL. 14, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus".
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035042; AAD56278.1; -.
 DR INTERPRO: IPR003006; -.
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match 63.4%; Score 59; DB 4; Length 113;
 Best Local Similarity 68.4%; Pred. No. 0.034;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAEAGGVVOPGXSDRL 20
 DB 2 VOLVESGGVOPGSLRL 20
 :||| ||||| ||||| |||

RESULT 5
 ID Q9UL93 PRELIMINARY; PRT; 116 AA.
 AC Q9UL93;
 DT 01-MAY-2000 (TREMBLERL. 13, Created)
 DT 01-MAY-2000 (TREMBLERL. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLERL. 14, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 GN VH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus".
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035042; AAD56278.1; -.
 DR INTERPRO: IPR003006; -.
 DR PFAM: PF00047; 19; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 116 AA; 12434 MW; ODA0348154DD6061 CRC64;

Query Match 63.4%; Score 59; DB 4; Length 116;
 Best Local Similarity 68.4%; Pred. No. 0.035;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAEAGGVVOPGXSDRL 20
 DB 1 VOLVESGGVOPGSLRL 19
 :||| ||||| ||||| |||

RESULT 6
 ID Q9UL71 PRELIMINARY; PRT; 121 AA.
 AC Q9UL71;
 Q9UL71;

AC 09UL71:
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035043; AAD56279.1; -
 DR HSSP; P01772; 2F84.
 DR INTERPRO: IPR003006; -
 DR PFAM; PF00047; 19; 1.
 FT NON_TER 1
 FT NON_TER 121
 SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFASD50736 CRC64;

Query Match 63.4%; Score 59; DB 4; Length 121;
 Best Local Similarity 68.4%; Pred. No. 0.037;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGGLVOPGXSURL 20
 Db 2 VOLVESGGVOPGSLURL 20

RESULT 7
 09UL84
 ID 09UL84 PRELIMINARY; PRT; 122 AA.
 AC 09UL84;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035043; AAD56279.1; -
 DR HSSP; P01772; 2F84.
 DR INTERPRO: IPR003006; -
 DR PFAM; PF00047; 19; 1.
 FT NON_TER 1
 FT NON_TER 122
 SQ SEQUENCE 122 AA; 13579 MW; 36054DA1366545B8 CRC64;

Query Match 63.4%; Score 59; DB 4; Length 122;
 Best Local Similarity 68.4%; Pred. No. 0.037;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGGLVOPGXSURL 20
 Db 2 VOLVESGGVOPGSLURL 20

RESULT 8
 09UL88
 ID 09UL88 PRELIMINARY; PRT; 131 AA.
 AC 09UL88;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT)
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus."
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035026; AAD56262.1; -
 DR INTERPRO: IPR003006; -
 DR PFAM; PF00047; 19; 1.
 FT NON_TER 1
 FT NON_TER 131
 SQ SEQUENCE 131 AA; 14142 MW; 96E7D668E375DEA0 CRC64;

Query Match 62.4%; Score 58; DB 4; Length 131;
 Best Local Similarity 68.4%; Pred. No. 0.059;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGGLVOPGXSURL 20
 Db 2 VOLVESGGGLVKPGSLURL 20

RESULT 9
 09OV16
 ID 09OV16 PRELIMINARY; PRT; 15 AA.
 AC 09OV16;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE PROLACTIN-BINDING PROTEIN (FRAGMENT)
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=95094032; PubMed=8000909;
 RA Cohen H., Cohen O., Gagnon J.;
 RT "Serum prolactin-binding protein (PRL-BP) of human and rat are
 identified as IgG."
 RL C.R. Acad. Sci., III, Sci. Vie 317:293-298(1994).
 DR HSSP; P01789; IMCP.
 SQ SEQUENCE 15 AA; 1469 MW; 35ED2512FF3FA369 CRC64;

Query Match 58.1%; Score 54; DB 11; Length 15;
 Best Local Similarity 78.6%; Pred. No. 0.023;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 LQLAESGGLVOPG 15
 Db 2 VOLVESGGGLVOPG 15

RESULT 10
 09Y509
 ID 09Y509 PRELIMINARY; PRT; 147 AA.
 AC 09Y509;

DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE VH3 PROTEIN (FRAGMENT).
 GN VH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96071149; PubMed=7475288;
 RA Cao J., Vescio R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C.,
 RA Lichtenstein A.K., Berenson J.R.;
 RT "A CD10-positive subset of malignant cells is identified in multiple
 RT myeloma using PCR with patient-specific immunoglobulin gene primers";
 RL Leukemia 9:1948-1953(1995).
 DR EMBL: S80860; AAD1439.1; -.
 DR HSSP: P01772; 2P84.
 DR INTERPRO: IPR003006; -.
 DR PFM: PF00047; 19; 1.
 FT NON_TER
 SO SEQUENCE 147 AA; 15768 MW; 8489FCAA7BC925C CRC64;

Query Match 58.1%; Score 54; DB 4; Length 147;
 Best Local Similarity 63.2%; Pred. No. 0.3;
 Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 LQAESGGVVOGXSRL 20
 DB 2 VLVESGGVVOGXSRL 20

RESULT 11

O9R1A4 PRELIMINARY; PRT; 437 AA.
 ID O9R1A4
 AC O9R1A4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE GAMMA1 HEAVY CHAIN OF MAB7 (FRAGMENT).
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
 RT "Cloning of CDNA's encoding for anti-white pine blister rust monoclonal
 RT antibody (Mab 7, its light and heavy chains) and construction of a
 RT single chain antibody (scFv).";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF152372; AAD40243.1; -.
 DR HSSP: P01842; 7FAB.
 DR INTERPRO: IPR003006; -.
 DR PFM: PF00047; 19; 4.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER
 FT NON_TER
 SO SEQUENCE 437 AA; 48142 MW; 5C3A7BB3E7D697C CRC64;

Query Match 58.1%; Score 54; DB 11; Length 437;
 Best Local Similarity 63.2%; Pred. No. 1;
 Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAESGGVVOGXSRL 20
 DB 1 VLVESGGVVOGXSRL 19

RESULT 12

O9QYF0 PRELIMINARY; PRT; 298 AA.
 ID O9QYF0
 AC O9QYF0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE CN 8 SCFV.
 GN CN 8.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-BALB/C; TISSUE-SPLEEN;
 RA Shinozaki N., Demura T., Fukuda H.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-BALB/C; TISSUE-SPLEEN;
 RA Shinozaki N., Demura T., Fukuda H.;
 RT "Isolation of a novel type of vascular cell wall-specific monoclonal
 RT antibody recognizing a cell polarity using a phage display subtraction
 RT method";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB036341; BAA8633.1; -.
 DR HSSP: P01607; 1REI.
 DR INTERPRO: IPR003006; -.
 DR PFM: PF00047; 19; 2.
 SO SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 50.5%; Score 47; DB 11; Length 298;
 Best Local Similarity 52.6%; Pred. No. 8.9;
 Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQAESGGVVOGXSRL 20
 DB 41 VKLOSGGVLVPGGSLK 59

RESULT 13

O9X7E8 PRELIMINARY; PRT; 308 AA.
 ID O9X7E8
 AC O9X7E8;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE PURATIVE PSEUDOURIDINE SYNTHASE.
 GN RUC.
 OS Mycobacterium leprae.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1769;
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93188700; PubMed=8446027;
 RA Seeger K.J., Harris D.;
 RT "Use of an ordered cosmid library to deduce the genomic organization
 RT of Mycobacterium leprae";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93188700; PubMed=8446027;
 RA Eigmeier K., Honore N., Woods S.A., Gaudron B., Cole S.T.;
 RT "Use of an ordered cosmid library to deduce the genomic organization
 RT of Mycobacterium leprae";
 RL Mol. Microbiol. 7:197-206(1993).
 DR EMBL: AL049478; CAB39580.1; -.

DR INTERPRO: IPR000613; -
 DR INTERPRO: IPR002990; -
 DR PFAM: PF00849; Pseudou_synth_2; 1.
 DR PROSITE: PS01129; PSI_RL0; 1.
 SQ SEQUENCE 308 AA; 33157 MW; 55C9AA34B81FAE59 CRC64;

Query Match 50.0%; Score 46.5; DB 2; Length 308;
 Best Local Similarity 65.0%; Pred.No. 11;
 Matches 13; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

OY 4 LAESGGVL--VQPGXSRL 20
 ||| ||| ||| ||| |||
 DB 35 LAEDGVELDGVQAGKSDRL 54

RESULT 14
 O9UC53 PRELIMINARY; PRT; 16 AA.
 AC O9UC53;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE 77 KDA SPONTANEOUS RECURRENT ABORTION-ASSOCIATED HUMAN EMBRYONIC
 DE ANTIGEN/IGWHII HOMOLOG (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96033130; PubMed=8582963;
 RA Shirahishi Y., Shirahishi Y., Yamamoto D., Hasegawa T., Kitamura W.,
 RA Miki S., Tanaka T., Suzuki T., Soma H.;
 RT "Diagnostic relevance of abortion-associated human embryonic antigen
 RT expressed on the cell surface of tumour promoter-treated Bloom
 RT syndrome cells."
 RL Hum. Reprod. 10:1694-1701(1995).
 SQ SEQUENCE 16 AA; 1626 MW; C9C5ED2512FF3FB9 CRC64;

Query Match 48.4%; Score 45; DB 4; Length 16;
 Best Local Similarity 71.4%; Pred.No. 0.72;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 LQLAESGGVLVQPG 15
 :||| ||| ||| |||
 DB 2 VOLVESGVLVQPG 15

RESULT 15
 Q08693 PRELIMINARY; PRT; 809 AA.
 ID Q08693;
 AC Q08693;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CHROMOSOME XV READING FRAME ORF YOR256C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97298311; PubMed=9153759;
 RA Jauniaux J.C., Polirey R.;
 RT "Sequencing analysis of a 36.8 kb fragment of yeast chromosome XV
 RT reveals 26 open reading frames including SEC63, CDC31, SUG2, GCD1,
 RT RBL2, PWT1, PAC1 and VPH1."

RL Yeast 13:483-487(1997).
 DR EMBL: 275164; CAA99478.1; -
 DR INTERPRO: IPR003137; -
 DR PFAM: PF02225; PA; 1.
 SQ SEQUENCE 809 AA; 91995 MW; D7D68C0A8C50ECB2 CRC64;

Query Match 48.4%; Score 45; DB 3; Length 809;
 Best Local Similarity 52.6%; Pred.No. 58;
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 MLQLAESGGVLVQPGXSDR 19
 :| ||| ||| ||| |||
 DB 380 LLSRLSSGGVTYDDGNSDR 398

Search completed: June 13, 2001, 14:29:44
 Job time: 545 sec

RESULT 2
US-08-545-809A-125
Sequence 125, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 118 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-125

Query Match 71.0%; Score 66; DB 3; Length 118;
Best Local Similarity 73.7%; Pred. No. 0.0012; Mismatches 14; Conservative 2; Indels 3; Gaps 0;

QY 2 LQLAESGGLVOPGXSDRL 20
DB 21 VOLVESGGVVPDGSRL 39

RESULT 3
US-07-977-696C-75
Sequence 75, Application US/07977696C
Patent No. 5792852
GENERAL INFORMATION:
APPLICANT: do Couto, Fernando J.R.
APPLICANT: Ceriani Dr., Roberto L.
APPLICANT: Peterson Dr., Jerry A.
APPLICANT: Padlan Dr., Eduardo A.
TITLE OF INVENTION: Analogue Peptides with Specificity
TITLE OF INVENTION: for Carcinomas and Kit and Diagnostic Vaccination
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLANSKI
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles

STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,696C
FILING DATE: 11-16-92
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Amzel Ph.D., Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38227
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 748-6868
TELEFAX: (510) 748-6688
TELEX: n.a.
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-977-696C-75

Query Match 67.7%; Score 63; DB 1; Length 30;
Best Local Similarity 73.7%; Pred. No. 0.00083; Mismatches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVOPGXSDRL 20
DB 2 VOLVESGGVVPDGSMDRL 20

RESULT 4
US-08-129-930B-75
Sequence 75, Application US/08129930B
Patent No. 5804187
GENERAL INFORMATION:
APPLICANT: do Couto Dr., Fernando J.R.
APPLICANT: Ceriani Dr., Roberto L.
APPLICANT: Peterson Dr., Jerry A.
APPLICANT: Padlan Dr., Eduardo A.
TITLE OF INVENTION: Analogue Peptides with Broad
TITLE OF INVENTION: Carcinoma Specificity, and Kit and
TITLE OF INVENTION: Diagnostic Vaccination and
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: V. AMZEL & ASSOC.
STREET: 2055 No. 5804187th Broadway, Suite 201
CITY: Walnut Creek
STATE: California
COUNTRY: USA
ZIP: 94596
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/129,930B
FILING DATE: September 30, 1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Amzel Ph.D., Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: CRFC-008A
TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 521-1333
TELEFAX: (510) 521-3541
TELEX: n.a.
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-129-930B-75

Query Match 67.7%; Score 63; DB 1; Length 30;
Best Local Similarity 73.7%; Pred. No. 0.00083;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 2 VOLVESGGVLVOPGSGMRL 20

RESULT 5
US-08-859-931A-2
Sequence 2, Application US/08859931A
Patent No. 5945510
GENERAL INFORMATION:
APPLICANT: FASANO, Alessio
TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A
TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
CITY: Washington, D.C.
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,931A
FILING DATE: 21 MAY 1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: KIT, Gordon
REGISTRATION NUMBER: 30,764
REFERENCE/DOCKET NUMBER: A-6901
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOHETICAL: NO
US-08-859-931A-2

Query Match 66.7%; Score 62; DB 2; Length 20;
Best Local Similarity 73.7%; Pred. No. 0.00077;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 2 VOLVESGGVLVOPGSGRL 20

RESULT 6
US-08-471-780C-80
Sequence 80, Application US/08471780C
Patent No. 5759808
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,780C
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA: FR 93401310.3

APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potler, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius

US-08-471-780C-80

Query Match 66.7%; Score 62; DB 1; Length 26;
Best Local Similarity 73.7%; Pred. No. 0.001;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 LQLAESGVLVOPGXSRL 20
DB 2 VOLVESGGVLVOPGSGRL 20

RESULT 7
US-08-467-282B-80
Sequence 80, Application US/08467282B
Patent No. 5800988
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.

US-08-467-282B-80
Sequence 80, Application US/08467282B
Patent No. 5800988
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
APPLICANT: Hamers, Raymond
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.

```

1 CLASSIFICATION: 536
2 PRIOR APPLICATION NUMBER:
3 APPLICATION NUMBER: US 08/106,944
4 FILING DATE: 17-AUG-1993
5 APPLICATION NUMBER: FR 92402326.0
6 FILING DATE: 21-AUG-1992
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: FR 93401310.3
9 FILING DATE: 21-MAY-1993
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Potter, Jane E.R.
12 REGISTRATION NUMBER: 33,332
13 REFERENCE/DOCKET NUMBER: 04958.0008-00000
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 202-408-4000
16 TELEFAX: 202-408-4400
17 INFORMATION FOR SEQ ID NO: 80:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 26 amino acids
20 TYPE: amino acid
21 STRANDEDNESS: single
22 TOPOLOGY: linear
23 MOLECULE TYPE: protein
24 ORIGINAL SOURCE:
25 ORGANISM: Camelus dromedarius
26 US-08-471-282A-80
27
28 Query Match 66.7%; Score 62; DB 2; Length 26;
29 Best Local Similarity 73.7%; Pred. No. 0.001;
30 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0.
31
32 QY 2 LQALSGGLVQPGXSRL 20
33 : 11 1111 1111 11
34 Db 2 VQLVESGGGLVQPGGSLRL 20
35
36 RESULT 9
37 US-08-466-710C-80
38 Sequence 80, Application US/08466710C
39 Patent No. 5874541
40 GENERAL INFORMATION:
41 APPLICANT: Casteleman, Cecile
42 APPLICANT: Hamers, Raymond
43 TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
44 NUMBER OF SEQUENCES: 130
45 CORRESPONDENCE ADDRESS:
46 ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner
47 STREET: 1300 I Street, N.W.
48 CITY: Washington
49 STATE: D.C.
50 COUNTRY: USA
51 ZIP: 20005-3315
52 COMPUTER READABLE FORM:
53 MEDIUM TYPE: floppy disk
54 COMPUTER: IBM PC compatible
55 OPERATING SYSTEM: PC-DOS/MS-DOS
56 SOFTWARE: Patentln Release #1.0, Version #1.25
57 CURRENT APPLICATION DATA:
58 APPLICATION NUMBER: US/08/466,710C
59 FILING DATE:
60 CLASSIFICATION:
61 PRIOR APPLICATION DATA:
62 APPLICATION NUMBER: US/08/106,944
63 FILING DATE: 17-AUG-1993
64 APPLICATION NUMBER: FR 92402326.0
65 FILING DATE: 21-AUG-1992
66 PRIOR APPLICATION DATA:
67 APPLICATION NUMBER: FR 93401310.3
68 FILING DATE: 21-MAY-1993
69 ATTORNEY/AGENT INFORMATION:
70 NAME: Potter, Jane E.R.
71 REGISTRATION NUMBER: 33,332
72

```

```

REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-466-710C-80

Query Match          66.7%  Score 62;  DB 2;  Length 26;
Best Local Similarity 73.7%  Pred. No. 0.001;
Matches 14;  Conservative 1;  Mismatches 4;  Indels 0;  Gaps 0

QY      2 LQLAESGVLVQPGXSRL 20
        :| | | | | | | | | |
Db       2 VLVESGGVLVQPGGSLRL 20

RESULT 10
US-08-468-739C-80
: Sequence 80, Application US/08468739C
: Patent No. 6015695
: GENERAL INFORMATION:
: APPLICANT: Casterman, Cecile
: APPLICANT: Hamers, Raymond
: TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
: NUMBER OF SEQUENCES: 130
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flinegan, Henderson, Farabow, Garrett & Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,739C
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/106,944
: FILING DATE: 17-AUG-1993
: APPLICATION NUMBER: FR 92402326.0
: FILING DATE: 21-AUG-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 93401310.3
: FILING DATE: 21-MAY-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Poter, Jane E.R.
: REGISTRATION NUMBER: 33,332
: REFERENCE/DOCKET NUMBER: 04958.0008-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 80:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 26 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
:

```

```

: ORGANISM: Camelus dromedarius
US-08-468-739C-80

Query Match          66.7%; Score 62; DB 3; Length 26;
Best Local Similarity 73.7%; Pred. No. 0.001;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY      2 LQLAESGGVLVQPGXSRL 20
        :|| |||| |||| | ||
        2 VOLVESGGVLVQPGGSLRL 20

RESULT 11
US-08-470-139-17
: Sequence 17, Application US/08470139
: Patent No. 5998586
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Interleukin-5 specific recombinant antibodies
: NUMBER OF SEQUENCES: 28
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/470,139
: FILING DATE: 06 JUNE-1995
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: TRUJILLO, DOREEN YATRO
: REGISTRATION NUMBER: 35,719
: REFERENCE/DOCKET NUMBER: CARP-0044
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 30 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-470-139-17

Query Match          66.7%; Score 62; DB 2; Length 30;
Best Local Similarity 73.7%; Pred. No. 0.0012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY      2 LQLAESGGVLVQPGXSRL 20
        :|| |||| |||| | ||
        2 VOLVESGGVLVQPGGSLRL 20

RESULT 12
US-08-211-202-116
: Sequence 116, Application US/08211202
: Patent No. 5565332
: GENERAL INFORMATION:
: APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
: APPLICANT: BAIER, Michael
: APPLICANT: JESPERS, Laurent Stephane Anne Therese
: APPLICANT: WINTER, Gregory Paul
: TITLE OF INVENTION: Production of chimeric antibodies - a
: TITLE OF INVENTION: combinatorial approach
: NUMBER OF SEQUENCES: 144
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
: ADDRESSEE: Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606-6402

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120252.3
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120377.8
FILING DATE: 25-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-211-202-116

Query Match 66.7%; Score 62; DB 1; Length 98;
Best Local Similarity 73.7%; Pred. No. 0.0043;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGXSRL 20
DB 2 VQLVESGGLVQPGXSRL 20

RESULT 13
US-07-942-245-37
Sequence 37, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUTLD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation

OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-37

Query Match 66.7%; Score 62; DB 1; Length 98;
Best Local Similarity 73.7%; Pred. No. 0.0043;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGXSRL 20
DB 2 VQLVESGGLVQPGXSRL 20

RESULT 14
US-08-665-202-31
Sequence 31, Application US/08665202
Patent No. 5977322
GENERAL INFORMATION:
APPLICANT: Marks, James D.
APPLICANT: Schier, Robert
TITLE OF INVENTION: No. 5977322e1 High Affinity Human Antibodies to
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,202
FILING DATE: 13-JUN-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307E-061410
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-665-202-31

Query Match 66.7%; Score 62; DB 2; Length 98;
Best Local Similarity 73.7%; Pred. No. 0.0043;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGASDRL 20
:|||||
DB 2 VQLVESGGGLVQPGGSLRL 20

RESULT 15
US-08-428-197-4
; Sequence 4, Application US/08428197
; Patent No. 5891438

GENERAL INFORMATION:

APPLICANT: SILVERMAN, GREGG J.
TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East - Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,197
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10555
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:

NAME: Howells, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: FD-2630
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

CLONE: SFL

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..108

US-08-428-197-4

Query Match 66.7%; Score 62; DB 2; Length 108;
Best Local Similarity 73.7%; Pred. No. 0.0048;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LQLAESGGLVQPGASDRL 20
:|||||
DB 2 VQLVESGGGLVQPGGSLRL 20

Search completed: June 13, 2001, 14:27:07
Job time: 628 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:46 ; Search time 150.28 Seconds
(without alignments)
4.184 Million cell updates/sec

Title: PCT-US01-05825A-28

Perfect score: 48

Sequence: 1 EVOLVESGXL 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A.Geneseq_0401.*

1: /SID56/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	97.9	100	21	B40083
2	47	97.9	119	16	R66308
3	47	97.9	150	20	W94492
4	46	95.8	11	20	W94492
5	46	95.8	11	21	W79134
6	46	95.8	18	16	W39953
7	46	95.8	18	16	W68167
8	46	95.8	20	19	W68169
9	46	95.8	20	20	W94487
10	46	95.8	20	21	Y84660
11	46	95.8	20	21	Y79130

12	46	95.8	30	17	R87049	Human group III b
13	46	95.8	68	21	B53646	Human colon cancer
14	46	95.8	85	21	Y64735	Human 5' EST relat
15	46	95.8	96	21	Y56646	Partial peptide fr
16	46	95.8	97	21	B40112	Anti-HIL12 antibod
17	46	95.8	97	21	B40132	Anti-HIL12 antibod
18	46	95.8	98	14	R34279	Human TNF binding
19	46	95.8	98	15	R52066	Heavy chain variab
20	46	95.8	98	16	R72074	DP54 VH region. H
21	46	95.8	98	19	W59614	Anti-RSV F protein
22	46	95.8	98	21	B40072	Anti-HIL12 antibod
23	46	95.8	98	21	B40084	Anti-HIL12 antibod
24	46	95.8	98	21	B40085	Anti-HIL12 antibod
25	46	95.8	98	21	B40086	Anti-HIL12 antibod
26	46	95.8	98	21	B40087	Anti-HIL12 antibod
27	46	95.8	98	21	B40094	Anti-HIL12 antibod
28	46	95.8	98	21	B40095	Anti-HIL12 antibod
29	46	95.8	98	21	B40096	Anti-HIL12 antibod
30	46	95.8	98	21	B40113	Anti-HIL12 antibod
31	46	95.8	98	21	B40130	Anti-HIL12 antibod
32	46	95.8	98	21	B40131	Anti-HIL12 antibod
33	46	95.8	98	21	B40133	Anti-HIL12 antibod
34	46	95.8	98	21	B40134	Anti-HIL12 antibod
35	46	95.8	98	21	B40136	Anti-HIL12 antibod
36	46	95.8	98	21	B40137	Anti-HIL12 antibod
37	46	95.8	98	21	B40138	Anti-HIL12 antibod
38	46	95.8	98	21	B40139	Anti-HIL12 antibod
39	46	95.8	98	21	B40140	Anti-HIL12 antibod
40	46	95.8	98	21	Y50964	Human FVIII antibo
41	46	95.8	98	21	Y56660	Partial peptide fr
42	46	95.8	98	22	B48022	Heavy chain sequen
43	46	95.8	99	21	Y50960	Human FVIII antibo
44	46	95.8	100	17	R6105	VH-15 autoantibod
45	46	95.8	100	17	R6106	VH-15 autoantibod

ALIGNMENTS

RESULT 1	
ID B40083	standard; Protein: 100 AA.
XX	
AC B40083;	
XX	
DT 05-FEB-2001	(first entry)
XX	
DE	Anti-hil12 antibody H chain V region amino acid sequence SEQ ID 609.
XX	
KW	Human; neutralising antibody; interleukin-12; IL-12; antiinflammatory;
KW	complementarily determining region; CDR; antirheumatic; antiarthritic;
KW	antisclerotic; neuroprotective; antiporotic; antiasthmatic; cardiant;
KW	antiparasitic; antibacterial; immunosuppressive; Crohn's disease;
KW	multiple sclerosis; rheumatoid arthritis.
XX	
OS	Homo sapiens.
XX	
PN	WO200056772-A1.
XX	
PD	28-SEP-2000.
XX	
PF	24-MAR-2000; 2000MO-US07946.
XX	
PR	25-MAR-1999; 99US-0126603.
XX	
PA	(BADI) BASF AG.
PA	(GEMV) GENETICS INST INC.
XX	
PI	Salfield JG, Roguska M, Paskind M, Banerjee S, Tracey DE, White M;
PI	Kaymakcalan Z, Labkovsky B, Sakorafas P, Friederich S, Myles A;
PI	Velman G, Venturini A, Warne NW, Widow A, Elyin JG, Duncan AR;
PI	Derlyshire EJ, Carmen S, Smith S, Hollet TL, Du Fou SL;
XX	

DR WPI: 2000-638250/61.

XX New human antibody specific for human interleukin-12 (IL-12) used to
PT treat disorders characterized by aberrant IL-12 expression e.g. Crohn's
PT disease and multiple sclerosis .

PS Claim 75: Page 121: 377pp: English.

XX This invention relates to a new human antibody specific for human
CC interleukin-12 (IL-12). The invention also includes antigen binding
CC portions that bind to IL-12. Sequences B39485-B39516 represent human
CC anti-IL-12 antibody heavy and light chain complementarity determining
CC region (CDR) amino acid sequences, and also includes variable region
CC amino acid sequences. Other variable region amino acid sequences are
CC given in B39517-B39560 and B40068-B40149. Sequences B39561-B39771
CC represent anti-IL-12 CDR3 related amino acid sequences, B39772-B40063
CC represent other CDR sequences. Light chain CDR3 consensus sequences are
CC given in B40064-B40067. Primers used in the identification and
CC construction of the antibodies of the invention are given in
CC C61062-C61071. The antibody of the invention is a neutralizing antibody
CC and has antirheumatic; antiallergic; antisclerotic; antiinflammatory;
CC neuroprotective; antipsoriatic; antistimatic; cardiant; antiparasitic;
CC antibacterial and immunosuppressive activity. The antibodies or
CC antigen-binding fragments are useful in the treatment of disorders
CC associated with detrimental release of human IL-12, especially Crohn's
CC disease, multiple sclerosis and rheumatoid arthritis. They can also be
CC used in the manufacture of a pharmaceutical composition to treat human
CC IL-12 disorders.

SO Sequence 100 AA:

Query Match 97.9%; Score 47; DB 21; Length 100;
Best Local Similarity 90.9%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVQLVESGXL 11
Db 1 evqlvesgsl 11
|||||
11

RESULT 2
R66308 R66308 standard; Protein; 119 AA.

AC R66308;
XX
XX 03-AUG-1995 (first entry)

DE Human immunoglobulin variable heavy chain #14.
XX
XX

KW Primer: PCR; amplify; human; immunoglobulin; variable; heavy chain;
KW cosmid; placenta; vector; pJB81; E.coli; mammalian.

OS Homo sapiens.

XX WO9426895-A.

XX 24-NOV-1994.

XX 10-MAY-1993; 93WO-JP00603.

XX 10-MAY-1993; 93WO-JP00603.

XX (NISR) JAPAN TOBACCO INC.

XX Honjo T, Matsuda F;

XX WPI: 1995-006791/01.

XX N-PSDB: Q78953.

PT DNA fragment comprising human immunoglobulin Vh genes - for the
PT production of human immunoglobulin in mammalian hosts

XX Claim 24: Page 49-50; 130pp: Japanese.

XX Protein sequences (R66295-51) are novel human immunoglobulin heavy chain
CC sequences encoded by novel isolated genes. The genes (O78939-79002) were
CC isolated and cloned from a series of cosmid constructs: Y202; Y103; Y21;
CC Y6; Y24; 3-31; M84; M18 and M31, by PCR amplification using primers
CC O78917-38. The genes are subdivided into 5 families of Vh genes. The
CC high molecular weight DNA from human placenta. The DNA was partially
CC digested with RsaI restriction enzyme. The fragments were separated by
CC gel electrophoresis and 35-45 kb fractions were collected. The fragments
CC were ligated with ClaI-digested cosmid vector pJB81. The ligation
CC products were then subcloned by colony hybridization. The Vh genes and
CC the DNA fragments encoding them are useful in producing human
CC immunoglobulin in mammalian hosts.

SO Sequence 119 AA:

Query Match 97.9%; Score 47; DB 16; Length 119;
Best Local Similarity 90.9%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVQLVESGXL 11
Db 20 evqlvesgsl 30
|||||
20

RESULT 3
Y07484 Y07484 standard; Protein; 150 AA.

AC Y07484;
XX
XX 17-AUG-1999 (first entry)

DE Anti-HIV-1 gp120 antibody 447-D VH chain protein.
XX
XX

KW Heavy chain; variable region; human; HIV-1; gp120; monoclonal antibody;
KW epitope; V3 loop; heterohybridoma; human immunodeficiency virus-1;
KW peripheral blood lymphocyte; Epstein-Barr virus; EBV; AIDS.

OS Homo sapiens.

XX US5914109-A.

XX 22-JUN-1999.

XX 21-NOV-1994; 94US-0345321.

XX 23-APR-1992; 92US-0872675.

XX 15-JUN-1990; 90US-0538451.

XX 12-APR-1991; 91US-0684090.

XX 21-NOV-1994; 94US-0345321.

XX (UTNY) UNIV NEW YORK STATE.

XX Gorny MK, Zolla-Pazner S;

XX WPI: 1999-370481/31.

XX N-PSDB: X79205.

PT Heterohybridoma producing human monoclonal antibodies to human
PT immunodeficiency virus-1

PS Claim 5; Fig 11; 42pp: English.

CC This sequence represents the heavy chain variable region of the human
CC anti-HIV-1 gp120 monoclonal antibody 447-D. The antibody is targeted
CC to an epitope on the V3 loop of gp120. The invention relates to the
CC generation of heterohybridomas producing human monoclonal antibodies

CC to a neutralising epitope of human immunodeficiency virus-1 (HIV-1)
CC prepared by transforming peripheral blood lymphocytes with Epstein-Barr
CC virus. The antibodies can be used to treat someone infected with HIV-1
CC or suffering from AIDS.
XX
XX Sequence 150 AA;
90

Query Match	97.98	Score 47	DB 20	Length 150
Best Local Similarly	90.98	Pred. No. 0.15		
Matches 10	Conservative 0	Mismatches 1	Indels 0	Gaps 0

```
QY      1 EVQLVESGGXL 11
          ||||·||| |
Db      20 evqlvesgga 30
```

RESULT	4
W94492	
ID	W94492 standard; peptide; 11 AA.

DT 21-APR-1999 (first entry)

Human adult intestine zonulin peptide

KW Zonulin; mammalian tight junction; zonula occludens toxin; ZOT;
Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody;
KW

OS Homo sapiens.

	key	Location/Qualifiers
FH	Misc-difference	10
FT		/note="unspecified"
FT		

PN WO9852415-A1.

PD 26-NOV-1998.

28-APR-1998; 98WO-US07636.

PR 21-MAY-1997; 97US-0859931.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A.

DR WPI; 1999-070123/06.

PT New purified zonulin - which is capable of reversibly opening
PT mammalian tight junctions, used for enhancing the delivery of agents
PT across intestinal and nasal mucosa and blood brain barrier

PS Example 3; Fig 6; 64pp; English.

The present invention describes pure zonulin which has an apparent molecular weight of 47 kD, as determined by SDS-PAGE, which is recognised by both anti- τ -tau polyclonal antibody and by anti-zonula occludens toxin (ZO1) polyclonal antibody, and is capable of reversibly opening mammalian tight junctions. Zonulin proteins function as physiological modulators of mammalian tight junctions. They can be used for enhancing the absorption of therapeutic agents across tight junctions of intestinal and nasal mucosa and across tight junctions of the blood brain barrier. Zonulin can be used with agents such as drugs, e.g. lisdacarin, adenosine, dobutamine, dopamine, epinephrine, norepinephrine, phenoltamine, doxapram, alfenitall, dezocin, malbuphine, buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurine, mivacurium, succinylcholine, cytarabine, mitomycin doxorubicin, vincristine, vinblastine, methicillin, mezlocillin, piperacillin, cefoxitin, cefendicid, cefmetazole and aztreonam, a hormone e.g. testosterone, nandrolone, menotropins, insulin, urofollitropin, CC

CC Interferon-alpha, interferon-beta, interferon-gamma, interleukin-1
CC (IL-1), IL-2, IL-4, IL-8, polyvalent IgG, specific IgG, IgA, or IgM.
CC The proteins can also be used for the production of antibodies which can
CC be used to assay for zonulin in body tissue or fluids, or in affinity-
CC purification of zonulin. The present sequence represents a zonulin
CC peptide.

Query Match	95.8%	Score 46;	DB 20;	Length 11;
Best Local Similarity	100.0%	Pred. No.	0.016;	
Matches 11; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

QY	1	EVQLVESGGXL	11
Db	1	evqlvesgqx1	11

RESULT	5	/
Y79134		
ID	Y79134	standard; Peptide; 11 AA.

DT 05-JUN-2000 (first entry)

DE Human adult intestine zonulin N-terminal sequence

Zonulin; antagonist; zonula occludens toxin receptor
 human; blood-brain barrier; antiinflammatory;
 gastrointestinal inflammation; therapy.

OS Homo sapiens.

	Location/Qualifiers
Key	10
Misc-difference	
FT	/note= "unidentified residue"
FT	

PN WO200007609-A1

PD 17-FEB-2000

28-JUL-1999; 99WO-US16683.

03-AUG-1998; 98US-0127815.

AA
PA
(UYMA-) UNIV MARYLAND BALTIMORE.

PI Fasano A.,

DR WPI; 2000-205565/18

PT New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -

PS Example 3; Fig 6; 69pp; English.

CC The present sequence is that of the N-terminal region of adult
CC human intestinal zonulin. The N-terminal sequences of human adult
CC and fetal zonulins (see Y79130-36) were compared with Vibrio cholerae
CC zonula occludens toxin (ZOT) to identify a common motif thought
CC to be involved in receptor binding. Peptide antagonists (see
CC Y79105-29) based on this motif are useful as antiinflammatory
CC agents for treatment of gastrointestinal inflammation, and for
CC treatment of conditions associated with breakdown of the blood-brain
CC barrier.

AA	Sequence	11 AA;
5Q		

Query match 95.88; Score 46; DB 21; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
| | | | | | | | | |
Db 1 evqlvesgxl 11

RESULT 6
W73953
ID W73953 standard; peptide; 18 AA.
XX
AC W73953;

DT 26-APR-1999 (first entry)

DE IgG heavy chain fragment, IgG (VH) (18).

KW IgG heavy chain; immune complex removal; conglutinin; infection;
KW serum sickness; chronic immune complex nephritis; parasitic disease;
KW autoimmune disease; neoplastic disease; immune pathogenesis; therapy.

XX Synthetic.

OS MO9506254-A1.

PN 02-MAR-1995.

PD 24-AUG-1994; 94WO-US09407.

PE 24-AUG-1993; 93US-0111071.

PR (IMMU-) APPLIED IMMUNE SCI INC.

PI Lee YW, Odonoghue G, Okarma TB;

DR WPI; 1995-106945/14.

PT New biomedical device for binding immune complexes - comprising
PT conglutinin covalently bound to a solid phase material

PS Example 4; Fig 11; 58pp; English.

CC This sequence represents a fragment of the IgG heavy
CC chain. This sequence was isolated using the biomedical device of the
CC invention, which comprises: (a) a biocompatible support; and
CC (b) conglutinin which is covalently bound to the support. The device can
CC be used for removing immune complexes (IC) from a fluid. This can be used
CC for treating a mammal or for evaluating the status of a mammal. It can be
CC used for treating diseases such as serum sickness, chronic immune complex
CC nephritis, bacterial infections, viral infections, parasitic diseases,
CC autoimmune diseases or neoplastic diseases. The devices can also be used
CC for purification of IC and to study immune pathogenesis and antigens
CC associated with a disease state. The covalently immobilised conglutinin
CC can capture IC with high capacity and specificity and allows elution
CC under mild conditions. The devices are stable with no loss of IC binding
CC activity after repeated use.

XX Sequence 18 AA;

Query Match 95.8%; Score 46; DB 16; Length 18;
Best Local Similarity 90.9%; Pred. No. 0.028;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
| | | | | | | | | |
Db 1 evqlvesgxl 11

RESULT 7
W68167
ID W68167 standard; peptide; 18 AA.

XX
AC W68167;

DT 06-OCT-1998 (first entry)

DE Complex scuPA/suPAR fibrinolytic activity regulating peptide 2.

KW Thrombolytic activity; soluble urokinase plasminogen activator receptor;
KW single chain urokinase type plasminogen activator; thromboembolism;
KW scuPA; suPAR; IgG; regulator; fibrinolytic activity; fibrin clot;
KW myocardial infarction; cerebro-vascular event; pulmonary embolism;
KW deep vein thrombosis; immunoglobulin; human.

XX Synthetic.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Misc-difference 10 /note="unidentified"

PN WO9825641-A1.

PD 18-JUN-1998.

PE 09-DEC-1997; 97WO-II00402.

PR 09-DEC-1996; 96US-0032676.

PA (HADA-) HADASTI MEDICAL RES SERVICES & DEV.

PI Higazi AA;

DR WPI; 1998-348262/30.

PT New compositions with thrombolytic activity for, e.g. treatment of
PT thromboembolism - comprise complex of single chain urokinase type
PT plasminogen activator, scuPA, and soluble urokinase plasminogen
PT activator receptor, suPAR

PS Example 5; Page 24; 46pp; English.

CC This represents a human immunoglobulin G (IgG) derived peptide sequence
CC that has a stimulating effect on the fibrinolytic activity of the scuPA/
CC suPAR complex of the invention. This sequence has similarity to the human
CC Ig heavy chain, Ig V-III region (HV3r). The invention provides a
CC a thromboembolic disorder associated with the treatment and prevention of
CC The composition comprises, as the active ingredient, a complex of a
CC single chain urokinase type plasminogen activator (scuPA) and a soluble
CC urokinase plasminogen activator receptor (suPAR). The complex (scuPA/
CC suPAR) has thrombolytic activity under physiological conditions and in
CC the presence of IgG, or of at least 1 IgG-derived peptide, and induces
CC fibrinolysis of fibrin clots. The compositions and complex are useful
CC for the treatment or prevention of thromboembolic disorders associated
CC with the formation of fibrin clots, especially myocardial infarctions,
CC cerebro-vascular events, pulmonary embolism and deep vein thrombosis.

XX Sequence 18 AA;

Query Match 95.8%; Score 46; DB 19; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
| | | | | | | | | |
Db 1 evqlvesgxl 11

RESULT 8
W68169
ID W68169 standard; peptide; 20 AA.

AC W68169;
 XX 06-OCT-1998 (first entry)
 DT Human Igg heavy chain V-III region (HV3r) peptide.
 DE
 XX
 KW Thrombolytic activity; soluble urokinase plasminogen activator receptor;
 KW single chain urokinase type plasminogen activator; thromboembolism;
 KW scuPA; supAR; Igg; regulator; fibrinolytic activity; fibrin clot;
 KW myocardial infarction; cerebro-vascular event; pulmonary embolism;
 KW deep vein thrombosis; immunoglobulin; human.
 XX
 OS Homo sapiens.
 XX
 FN W09825641-A1.
 XX
 PD 18-JUN-1998.
 XX
 PF 09-DEC-1997; 97WO-IL00402.
 XX
 PR 09-DEC-1996; 96US-0032676.
 XX
 PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 XX
 PI H1ge21 AA;
 XX
 DR WPI: 1998-348262/30.
 XX
 PT New compositions with thrombolytic activity for, e.g. treatment of
 PT thromboembolism - comprise complex of single chain urokinase type
 PT plasminogen activator, scuPA, and soluble urokinase plasminogen
 PT activator receptor, supAR
 XX
 XX Example 5; Page 25; 46pp; English.
 PS
 CC This represents a peptide sequence from the human immunoglobulin G
 CC (IgG) heavy chain, IG V-III region (HV3r). An Igg derived peptide
 CC sequence having similarity to this sequence has a stimulating effect on
 CC the fibrinolytic activity of the scuPA/supAR complex of the invention.
 CC The invention provides a thrombolytic therapeutic composition for the
 CC treatment and prevention of a thromboembolic disorder associated with
 CC the formation of fibrin clots. The composition comprises, as the active
 CC ingredient, a complex of a single chain urokinase type plasminogen
 CC activator (scuPA) and a soluble urokinase plasminogen activator receptor
 CC (supAR). The complex (scuPA/supAR) has thrombolytic activity under
 CC physiological conditions and in the presence of Igg, or of at least 1
 CC Igg-derived peptide, and induces fibrinolysis of fibrin clots. The
 CC compositions and complex are useful for the treatment or prevention of
 CC thromboembolic disorders associated with the formation of fibrin clots,
 CC especially myocardial infarctions, cerebro-vascular events, pulmonary
 CC embolism and deep vein thrombosis.
 CC
 XX Sequence 20 AA:
 SQ
 Query Match 95.8%; Score 46; DB 19; Length 20;
 Best Local Similarity 90.9%; Pred. No. 0.032;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EVOLVESGXL 11
 |||||
 DB 1 evqlvesgdl 11
 RESULT 9
 W94487 ID W94487 standard; peptide: 20 AA.
 XX
 AC W94487;
 XX
 DT 21-APR-1999 (first entry)
 DE Human adult heart zonulin N-terminal peptide.
 XX

XX zonulin; mammalian tight junction; zonula occludens toxin; ZOT;
 KW Vibrio cholerae; vaccine; cholera toxin; polyclonal antibody;
 KW intestinal mucosa; nasal mucosa; blood brain barrier.
 XX
 OS Homo sapiens.
 XX
 FN W09852415-A1.
 XX
 PD 26-NOV-1998.
 XX
 PF 28-APR-1998; 98WO-US07636.
 XX
 PR 21-MAY-1997; 97US-0859931.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Fasano A;
 XX
 DR WPI: 1999-070123/06.
 XX
 PT New purified zonulin - which is capable of reversibly opening
 PT mammalian tight junctions, used for enhancing the delivery of agents
 PT across intestinal and nasal mucosa and blood brain barrier
 XX
 PS Claim 2; Page 45; 64pp; English.
 XX
 CC The present invention describes pure zonulin which has an apparent
 CC molecular weight of 47 kD, as determined by SDS-PAGE, which is
 CC recognised by both anti-tau polyclonal antibody and by anti-zonula
 CC occludens toxin (ZOT) polyclonal antibody, and is capable of reversibly
 CC opening mammalian tight junctions. Zonulin proteins function as
 CC physiological modulators of mammalian tight junctions. They can be used
 CC for enhancing the absorption of therapeutic agents across tight
 CC junctions of intestinal and nasal mucosa and across tight junctions of
 CC the blood brain barrier. Zonulin can be used with agents such as drugs,
 CC e.g. lidocaine, adenosine, dobutamine, dopamine, epinephrine,
 CC norepinephrine, phenolamine, doxapram, alfentanil, dezocin, nalbuphine,
 CC buprenorphine, naloxone, ketorolac, midazolam, propofol, metacurline,
 CC mycurline, succinylcholine, cytarabine, milomycin doxorubicin,
 CC vincristine, vinblastine, methicillin, mezlocillin, piperacillin,
 CC cefoxitin, cefenicol, cefmetazole and aztreonam, a hormone e.g.
 CC testosterone, nandrolone, menotropins, insulin, urofollitropin,
 CC interferon-alpha, interferon-beta, interferon-gamma, interleukin-1
 CC (IL-1), IL-2, IL-4, IL-8, polyvalent Igg, specific Igg, IgA, or IgM.
 CC The proteins can also be used for the production of antibodies which can
 CC be used to assay for zonulin in body tissue or fluids, or in affinity-
 CC purification of zonulin. The present sequence represents an N-terminal
 CC peptide of zonulin.
 CC
 XX Sequence 20 AA:
 SQ
 Query Match 95.8%; Score 46; DB 20; Length 20;
 Best Local Similarity 90.9%; Pred. No. 0.032;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EVOLVESGXL 11
 |||||
 DB 1 evqlvesggl 11
 RESULT 10
 Y84660 ID Y84660 standard; peptide: 20 AA.
 XX
 AC Y84660;
 XX
 DT 25-JUL-2000 (first entry)
 DE N-terminal sequence of a human zonulin protein of 47 kDa.
 XX
 DE Human; ZOT; zonula occludens toxin; zonulin; antigen presenting cell;
 XX

KW APC; lymphocyte proliferation; antigen; auto-immune disorder;
 KW immune-related disorder; immune system rejection; multiple sclerosis;
 KW organ transplantation; inflammatory disease; allergic disease;
 KW rheumatoid arthritis; insulin dependent diabetes mellitus;
 KW celiac disease; Sjogren's syndrome; systemic lupus erythematosus;
 KW auto-immune thyroiditis; idiopathic thrombocytopenic purpura;
 KW hemolytic anemia; Grave's disease; Addison disease; autoimmune orchitis;
 KW pernicious anemia; vasculitis; autoimmune coagulopathy; polymyositis;
 KW myasthenia gravis; polynuropathy; pemphigus; rheumatic carditis;
 KW dermatomyositis; scleroderma; asthma; psoriasis; eczematous dermatitis;
 KW Kaposi's sarcoma; inflammatory bowel disease; proliferative disorder.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 10 /note="any amino acid"
 FT
 PN W0200015252-A1.
 PD 23-MAR-2000.
 PF 09-SEP-1999; 99MO-US18842.
 PR 14-SEP-1998; 98US-0100266.
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA Fasano A, Szelein MB, Lu R, Tanner MK;
 PI WPI; 2000-271257/23.
 DR
 XX
 PT Suppression of antigen presenting cell mediated lymphocyte
 PT proliferation, by administering a Zot-related immunoregulator useful
 PT for treating immune-related disorders, immune system rejection
 PT subsequent to tissue or organ transplantation
 PS Disclosure; Page 26; 95pp; English.
 XX
 CC The present sequence represents the N-terminal of a human zonulin
 CC polypeptide. The specification describes a method of suppressing
 CC antigen presenting cell (APC)-mediated lymphocyte proliferation in a
 CC mammalian host pre-exposed to a particular antigen. The method comprises
 CC administering to the host an effective amount of a Zot-related
 CC immunoregulator selected from Zot (zonula occludens toxin) or zonulin,
 CC the amount effective to down-regulate the activity of the APC. The
 CC method can be used to down-regulate APC-mediated lymphocyte proliferation
 CC in mammalian hosts suffering from auto-immune or immune-related
 CC disorders, immune system rejection subsequent to tissue or organ
 CC transplantation, or inflammatory or allergic diseases. The auto-immune
 CC or immune related disorders include multiple sclerosis, rheumatoid
 CC arthritis, insulin dependent diabetes mellitus, celiac disease,
 CC Sjogren's syndrome, systemic lupus erythematosus, auto-immune
 CC thyroiditis, idiopathic thrombocytopenic purpura, hemolytic anemia,
 CC Grave's disease, Addison disease, autoimmune orchitis, pernicious
 CC anemia, vasculitis, autoimmune coagulopathy, myasthenia gravis,
 CC polynuropathy, pemphigus, rheumatic carditis, polymyositis,
 CC dermatomyositis, and scleroderma. The inflammatory or allergic disease
 CC or disorder is selected from asthma, psoriasis, eczematous dermatitis,
 CC Kaposi's sarcoma, multiple sclerosis, inflammatory bowel disease,
 CC proliferative disorders of smooth muscle cells, and inflammatory
 CC conditions associated with mycolic, viral, parasitic, or bacterial
 CC infections.
 CC
 SO Sequence 20 AA:
 QY
 Query Match 95.8%; Score 46; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 EVOLVESGXL 11
 |||||

Db 1 evqlvesgxl 11
 RESULT 11
 ID 179130 standard; Peptide; 20 AA.
 XX 179130:
 AC Y79130:
 DT 05-JUN-2000 (first entry)
 XX
 DE Human adult heart zonulin N-terminal sequence.
 KW Zonulin; antagonist; zonula occludens toxin receptor;
 KW human; blood-brain barrier; antiinflammatory;
 KW gastrointestinal inflammation; therapy.
 OS Homo sapiens.
 FH W0200007609-A1.
 PN 17-FEB-2000.
 PD 28-JUL-1999; 99MO-US16683.
 PF 03-AUG-1998; 98US-0127815.
 PR (UYMA-) UNIV MARYLAND BALTIMORE.
 PA Fasano A;
 PI WPI; 2000-205565/18.
 DR
 XX
 PT New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PT
 XX Example 3; Fig 6; 69pp; English.
 PS
 CC The present sequence is that of the N-terminal region of adult
 CC human heart zonulin. The N-terminal sequences of human adult and
 CC foetal zonulins (see Y79130-36) were compared with Vibrio cholerae
 CC zonula occludens toxin (ZOT) to identify a common motif thought
 CC to be involved in receptor binding. Peptide antagonists (see
 CC Y79105-29) based on this motif are useful as antiinflammatory
 CC agents for treatment of gastrointestinal inflammation, and for
 CC treatment of conditions associated with breakdown of the blood-brain
 CC barrier.
 CC
 SO Sequence 20 AA:
 QY
 Query Match 95.8%; Score 46; DB 21; Length 20;
 Best Local Similarity 90.9%; Pred. No. 0.032; 1; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1;
 Db 1 evqlvesgxl 11
 |||||

RESULT 12
 ID R87049 standard; Peptide; 30 AA.
 XX R87049:
 AC R87049:
 DT 25-JUN-1996 (first entry)
 XX
 DE Human group III heavy chain framework 1.
 KW Humanised antibody; Interleukin-5; IL-5; recombinant antibody;
 KW antibody engineering; monoclonal antibody; Mab; 39D10; CDR.

KW complementarity determining region; heavy chain; framework;
 KW eosinophilia; allergy; asthma.
 OS Homo sapiens.
 XX MO9535375-A1.
 XX
 PD 28-DEC-1995.
 XX
 PF 16-JUN-1995; 95WO-GB01411.
 XX
 PR 17-JUN-1994; 94GB-0012230.
 XX
 PA (CLLT) CELLTech THERAPEUTICS LTD.
 XX
 PI Athwal DS, Bodmer MW, Entage JS;
 XX
 DR WPI; 1996-058412/06.
 XX
 PT Anti-human IL-5 recombinant antibody - useful for preventing or
 PT reducing eosinophilia and for treating certain allergic diseases,
 PT esp. asthma
 XX
 PS Example 3; Fig 4; 69pp; English.
 XX
 CC Framework regions (R87049-52) of human group III (gp3) germ line
 CC antibody heavy chain showed homology to corresponding regions
 CC (R87053-56, respectively) of the rat anti-human Interleukin-5
 CC monoclonal antibody 39D10 heavy chain (see R87039). This homology
 CC was utilised in the prodn. of a humanised 39D10 VH (R87058) in
 CC which rat 39D10 VH complementarity determining regions were grafted
 CC into the human gp3 framework.
 CC
 SQ Sequence 30 AA:

Query Match 95.8%; Score 46; DB 17; Length 30;
 Best Local Similarity 90.9%; Pred. No. 0.049;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVQVESGGL 11
 |||||
 DB 1 evqlvesgggl 11

RESULT 13

B53646 B53646 standard; Protein: 68 AA.

AC B53646;

DT 09-MAR-2001 (first entry)

DE Human colon cancer antigen protein sequence SEQ ID NO:1186.

XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW Identification; cytostatic; cardioactive; neuroprotective; vulnery;
 KW immunomodulatory; muscular; gynaecological; gastrointestinal;
 KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
 KW neural disorder; immune system disorder; muscular disorder;
 KW reproductive disorder; gastrointestinal disorder; renal disorder;
 KW infectious disease; cardiovascular disorder.

XX Homo sapiens.

PN WO200055351-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05883.

PR 12-MAR-1999; 99US-0124270.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2000-587534/55.

DR N-PSDB; C98403.

PT Colon cancer associated gene sequences, referred to as colon cancer
 PT antigens, useful for the treatment, prevention, and diagnosis of colon
 PT disorders such as colon cancer -

PS Claim 11; Page 1766; 2104pp; English.

CC C97991 to C98763 encode the human colon cancer associated proteins,
 CC called human colon cancer antigens, given in B53234 to B54006. The human
 CC colon cancer antigens can have cytostatic, cardioactive, muscular;
 CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
 CC vulnerability, nephrotropic, antiinfective and antibacterial activities, and
 CC can be used in gene therapy. The colon cancer antigen polynucleotides,
 CC proteins and antibodies to the proteins are useful for the prevention,
 CC treatment and diagnosis of colon disorders, such as colon cancer. The
 CC polynucleotides may be used in diagnostics and research, such as for
 CC chromosome identification, and as hybridisation probes. The proteins
 CC may also be used to prevent diseases such as neural disorders, immune
 CC system disorders, muscular disorders, reproductive disorders, immune
 CC gastrointestinal disorders, wounds, renal disorders, infectious
 CC diseases, and cardiovascular disorders. C98764 to C98772 and B54007
 CC represent sequences used in the exemplification of the present
 CC invention.

SQ Sequence 68 AA:

Query Match 95.8%; Score 46; DB 21; Length 68;
 Best Local Similarity 90.9%; Pred. No. 0.12;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVQVESGGL 11
 |||||
 DB 39 evqlvesgggl 49

RESULT 14

Y64735 Y64735 standard; Protein: 85 AA.

AC Y64735;

DT 01-FEB-2000 (first entry)

DE Human 5' EST related polypeptide SEQ ID NO:896.

XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification.

OS Homo sapiens.

PN WO9953051-A2.

PD 21-OCT-1999.

PF 09-APR-1999; 99WO-IB00712.

PR 09-APR-1998; 98US-0057719.

PR 28-APR-1998; 98US-0069047.

PA (GEST) GENSET.

PI Dumas Mline Edwards J, Duclert A, Giordano J;

DR WPI; 2000-038446/03.

```

XX Claim 20: Page 54; 123pp; English.
PS
XX
CC The invention provides an antibody (Ab) comprising donor CDRs
CC (complementarity determining regions) derived from a non-human antigen-
CC specific donor antibody, and an acceptor framework from a non-human
CC primate. The Abs are prepared by grafting CDRs from a non-human antigen-
CC specific donor antibody onto homologous Old World ape or monkey acceptor
CC frameworks. The Abs have reduced immunogenicity and are better tolerated
CC in humans (because of the close similarity between the human and primate
CC proteins), but retain the full antigen-binding affinity of the donor
CC antibody.
XX
SQ Sequence 96 AA:
OY 1 EVOLVESGGL 11
   |||||
   |||||
   |||||
Db 1 EVGLVESGGL 11

```

```

Query Match      95.8%;   Score 46;   DB 21;   Length 96;
Best Local Similarity 90.9%;
Matches 10;   Conservative 0;   Mismatches 1;   Indels 0;
QY      1  EYQIVGSGCXL 11
        |||||
Db       1  evqlvesg9gl 11

```

```

Query Match      95.8%;   Score 46;   DB 21;   Length 96;
Best Local Similarity 90.9%;
Matches 10;   Conservative 0;   Mismatches 1;   Indels 0;
QY      1  EYQIVGSGCXL 11
        |||||
Db       1  evqlvesg9gl 11

```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:12 : Search time 87.97 Seconds
(without alignments)
8.593 Million cell updates/sec

Title: PCT-US01-05825A-28

Perfect score: 48

Sequence: 1 EVOLVESGXL 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.67:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	97.9	108	2	PH1010
2	47	97.9	116	1	AIH02P
3	46	95.8	12	2	S21205
4	46	95.8	97	2	S26890
5	46	95.8	97	2	S46462
6	46	95.8	98	1	HYMS96
7	46	95.8	98	2	PL0121
8	46	95.8	98	2	PL0123
9	46	95.8	98	2	S26896
10	46	95.8	98	2	S29545
11	46	95.8	98	2	S26937
12	46	95.8	98	2	S26932
13	46	95.8	98	2	S26891
14	46	95.8	98	2	S26940
15	46	95.8	98	2	S26894
16	46	95.8	98	2	S26933
17	46	95.8	98	2	S26934
18	46	95.8	100	2	PL0132
19	46	95.8	100	2	S69886
20	46	95.8	100	2	S26925
21	46	95.8	108	2	PH1015
22	46	95.8	108	2	PH1011
23	46	95.8	110	2	PH1014
24	46	95.8	112	2	S25572
25	46	95.8	113	2	JL0049
26	46	95.8	113	2	S25575
27	46	95.8	114	1	AVDGM
28	46	95.8	114	2	S36280
29	46	95.8	114	2	S36280

30	46	95.8	116	1	M3HUGL	Ig heavy chain V-I
31	46	95.8	116	2	S31668	Ig heavy chain V-I
32	46	95.8	116	2	S21979	Ig heavy chain V-I
33	46	95.8	116	2	S17080	Ig heavy chain V-g
34	46	95.8	116	2	B28966	Ig heavy chain pre
35	46	95.8	117	1	HYMS34	Ig heavy chain pre
36	46	95.8	117	2	S78486	Ig heavy chain V-I
37	46	95.8	117	2	S17079	Ig heavy chain V-g
38	46	95.8	117	2	A34964	Ig heavy chain pre
39	46	95.8	117	2	S36259	Ig heavy chain V-I
40	46	95.8	117	2	S34012	Ig heavy chain V-I
41	46	95.8	117	2	S21980	Ig heavy chain V-g
42	46	95.8	117	2	S31109	Ig heavy chain - h
43	46	95.8	118	2	S31121	Ig heavy chain - h
44	46	95.8	118	2	S69132	Ig heavy chain V-I
45	46	95.8	119	1	GIHUTE	Ig heavy chain V-I

ALIGNMENTS

RESULT 1

PH1010 Ig heavy chain V region (Clone 17s.93) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C:Accession: PH1010

R:Tillman, D.M.; You, N.T.; Hill, R.J.; Marton, T.N.

J. Exp. Med. 176, 761-779, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective

A:Reference number: PH0971; MUID:92381444

A:Accession: PH1010

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-108 <TIL>

A:Experimental source: B cell, strain [NZB x NZW]F1

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin

F:15-97/Domain: Immunoglobulin homology <IMM>

Query Match 97.9% Score 47; DB 2; Length 108;
Best Local Similarity 90.9% Pred. No. 0.036; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
DB 1 EVOLVESGXL 11

RESULT 2

AIH02P Ig heavy chain V-III region (Zap) - human (tentative sequence)

C:Species: Homo sapiens (man)

C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000

C:Accession: A02061

R:Capra, J.D.; Kenne, J.M.

Proc. Natl. Acad. Sci. U.S.A. 71, 845-848, 1974

A:Title: Variable region sequences of five human immunoglobulin heavy chains of the V

A:Reference number: A03794; MUID:74142702

A:Accession: A02061

A:Molecule type: protein

A:Residues: 1-116 <CAP>

C:Comment: This chain was isolated from an IgA1 myeloma protein.

C:Genetics:

A:Gene: GDB:IGHV1

A:Cross-references: GDB:128528; OMIM:147070

A:Map position: 14q32.33-14q32.33

C:Superfamily: Immunoglobulin V region; Immunoglobulin homology

C:Keywords: heterotetramer; Immunoglobulin

F:15-98/Domain: Immunoglobulin homology <IMM>

F:22-96/Disulfide bonds: #status predicted

Query Match 97.8%; Score 46; DB 1; Length 116;
 Best Local Similarity 90.9%; Pred. No. 0.039;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 3

S21205
 Ig heavy chain V region - human
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 16-Aug-1996
 C:Accession: S21205
 R:Maxly, R.; Stigbrand, T.
 Eur. J. Biochem. 205: 341-345, 1992
 A:Title: Placental alkaline phosphatase has a binding site for the human immunoglobulin
 A:Reference number: S21205; MUID:92209522
 A:Accession: S21205
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-112 <MAK>
 C:Keywords: heterotetramer; immunoglobulin

Query Match

S26890
 Ig heavy chain V region (DP-48) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
 C:Accession: S26890
 R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
 J. Mol. Biol. 227, 776-798, 1992
 A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
 A:Reference number: S26885; MUID:93021117
 A:Accession: S26890
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-97 <TOM>
 A:Cross-references: EMBL:Z12348; NID:932916; PIDN:CAA78218.1; PID:932917
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>

RESULT 4

S26890
 Ig heavy chain V region (DP-48) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
 C:Accession: S26890
 R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
 J. Mol. Biol. 227, 776-798, 1992
 A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
 A:Reference number: S26885; MUID:93021117
 A:Accession: S26890
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-97 <TOM>
 A:Cross-references: EMBL:Z12348; NID:932916; PIDN:CAA78218.1; PID:932917
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>

Query Match

S46462
 Ig heavy chain V region (YAC-5) - human
 C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
 C:Accession: S46462
 R:Cook, G.P.; Tomlinson, I.M.; Walter, G.; Rietman, H.; Carter, N.P.; Bulwela, L.; Wirt
 Nature Genet. 7, 162-168, 1994

Best Local Similarity 95.8%; Score 46; DB 2; Length 97;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 5

S46462
 Ig heavy chain V region (YAC-5) - human
 C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
 C:Accession: S46462
 R:Cook, G.P.; Tomlinson, I.M.; Walter, G.; Rietman, H.; Carter, N.P.; Bulwela, L.; Wirt
 Nature Genet. 7, 162-168, 1994

A:Title: A map of the human immunoglobulin V(H) locus completed by analysis of the te
 A:Reference number: S46460; MUID:95004581
 A:Accession: S46462
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-97 <COO>
 A:Cross-references: EMBL:Z27504; NID:9505430; PIDN:CAA81824.1; PID:9505431
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 95.8%; Score 46; DB 2; Length 97;
 Best Local Similarity 90.9%; Pred. No. 0.051;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 6

HVMS96
 Ig heavy chain V region (6-96) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 31-Mar-1997
 C:Accession: J70501
 R:Levy, N.S.; Malipiero, U.V.; Lebecque, S.G.; Gearhart, P.J.
 J. Exp. Med. 169, 2007-2019, 1989
 A:Title: Early onset of somatic mutation in immunoglobulin VH genes during the primar
 A:Reference number: J70501; MUID:89279149
 A:Accession: J70501
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-98 <LEV>
 A:Experimental source: strain BALB/cJ
 A:Note: This sequence belongs to the VH7183 subfamily
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-98/Domain: immunoglobulin homology <IMM>
 F:22-96/Disulfide bonds: #status predicted

Query Match 95.8%; Score 46; DB 1; Length 98;
 Best Local Similarity 90.9%; Pred. No. 0.052;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 7

PL0121
 Ig heavy chain V-III region (TD-Vp) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Aug-1996
 C:Accession: PL0121
 R:Bird, U.; Gallili, N.; Link, M.; Stiles, D.; Sklar, J.
 J. Exp. Med. 168, 229-245, 1988
 A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobu
 A:Reference number: PL0116; MUID:88286083
 A:Accession: PL0121
 A:Molecule type: mRNA
 A:Residues: 1-98 <BIR>
 A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL
 A:Note: the sequence shows the V region (TD-Vp) from one of five DNA rearrangements f
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin
 F:15-98/Domain: immunoglobulin homology <IMM>
 F:31-35/Region: complementarity-determining 1
 F:49-65/Region: complementarity-determining 2

Query Match 95.8%; Score 46; DB 2; Length 98;
Best Local Similarity 90.9%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
|||||||
DB 1 EVOLVESGGL 11

RESULT 8
Ig heavy chain V-III region (TD-Vr) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
C:Accession: P10123; S26897

R:Bird, J.; Gallili, N.; Link, M.; Stites, D.; Sklar, J.
J. Exp. Med. 168, 229-245, 1988

A:Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin A:Reference number: P10116; MUID:88286083

A:Accession: P10123

A:Molecule type: mRNA

A:Residues: 1-98 <BIR>

A:Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL

A:Note: the sequence shows the V region (TD-Vr) from a nonproductive DNA rearrangement

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.

J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V A:Reference number: S26885; MUID:93021117

A:Accession: S26897

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-98 <TOM>

A:Cross-references: EMBL:212354; NID:932930; PID:CAA78224.1; PID:932931

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>
F:31-35/Region: complementarity-determining 1
F:49-65/Region: complementarity-determining 2

Query Match 95.8%; Score 46; DB 2; Length 98;
Best Local Similarity 90.9%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
|||||||
DB 1 EVOLVESGGL 11

RESULT 9
Ig heavy chain V region (DP-53) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
C:Accession: S26896

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V A:Reference number: S26885; MUID:93021117

A:Accession: S26896

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-98 <TOM>

A:Cross-references: EMBL:212353

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 95.8%; Score 46; DB 2; Length 98;
Best Local Similarity 90.9%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
|||||||
DB 1 EVOLVESGGL 11

RESULT 10

Ig heavy chain V region (COS 6) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jan-1994 #sequence_revision 17-Nov-1995 #text_change 23-Jul-1999
C:Accession: S29545

R:Tomlinson, M.; Walter, G.; Cook, G.P.; Winter, G.
submitted to the EMBL Data Library, October 1992

A:Reference number: S29543

A:Accession: S29545

A:Molecule type: DNA

A:Residues: 1-98 <TOM>

A:Cross-references: EMBL:217392; NID:932840; PID:CAA78996.1; PID:932841

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

OY 1 EVOLVESGXL 11
|||||||
DB 1 EVOLVESGGL 11

RESULT 11

Ig heavy chain V region (DP-31) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26927

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of A:Reference number: S26885; MUID:93021117

A:Accession: S26927

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-98 <TOM>

A:Cross-references: EMBL:212333; NID:932885; PID:CAA78203.1; PID:932886

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

OY 1 EVOLVESGXL 11
|||||||
DB 1 EVOLVESGGL 11

RESULT 12

Ig heavy chain V region (DP-39) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26932

R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992

A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of A:Reference number: S26885; MUID:93021117

A:Accession: S26932

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12339; NID:q32898; PIDN:CAA78209.1; PID:q32899
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match
Best Local Similarity 95.8%; Score 46; DB 2; Length 98;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
|||||||
Db 1 EVOLVESGGL 11

RESULT 13
S26891
Ig heavy chain V region (DP-58) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26891
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12358; NID:q32935; PIDN:CAA78228.1; PID:q32936
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match
Best Local Similarity 95.8%; Score 46; DB 2; Length 98;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
|||||||
Db 1 EVOLVESGGL 11

RESULT 14
S26940
Ig heavy chain V region (DP-77) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 17-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26940
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of
A:Reference number: S26885; MUID:93021117
A:Accession: S26940
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z14073; NID:q32973; PIDN:CAA78453.1; PID:q32974
A:Note: the nucleotide sequence was submitted to the EMBL data library, July 1992
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match
Best Local Similarity 95.8%; Score 46; DB 2; Length 98;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGGL 11
|||||||

RESULT 15
S26894
Ig heavy chain V region (DP-51) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26894
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewelyn, M.B.; Winter, G.
J. Mol. Biol. 227, 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups o
A:Reference number: S26885; MUID:93021117
A:Accession: S26894
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12351; NID:q32924; PIDN:CAA78221.1; PID:q32925
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match
Best Local Similarity 95.8%; Score 46; DB 2; Length 98;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
|||||||
Db 1 EVOLVESGGL 11

Search completed: June 13, 2001, 14:23:12
Job time: 743 sec

This Page Blank (uspto)

RA Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
 RT "Early onset of somatic mutation in immunoglobulin VH genes during
 the primary immune response.";
 RL J. Exp. Med. 169:2007-2019(1989).
 CC -I- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
 DR PIR: J050501; HVMS96.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON_TER 98
 SQ SEQUENCE 98 AA; 11007 MW; B8644F7F92FBF95B CRC64;

Query Match
 Best Local Similarity 95.8%; Score 46; DB 1; Length 98;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 1 EVOLVESGGL 11

RESULT 3
 HV01_CANFA STANDARD; PRT; 114 AA.
 AC P01784;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V REGION GOM.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=77242268; Pubmed=407924;
 RA Wasserman R.L., Capra J.D.;
 RT "Primary structure of the variable regions of two canine
 immunoglobulin heavy chains.";
 RL Biochemistry 16:3160-3168(1977).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
 DR PIR: A02067; AVDGM.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON_TER 114
 SQ SEQUENCE 114 AA; 12430 MW; B1D4745D2C4E13C4 CRC64;

Query Match
 Best Local Similarity 95.8%; Score 46; DB 1; Length 114;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 1 EVOLVESGGL 11

RESULT 4
 HV05_CARAU STANDARD; PRT; 116 AA.
 AC P19181;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V REGION 5A PRECURSOR.
 OS Carassius auratus (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=88144476; Pubmed=3125551;
 RA Wilson M.R., Middleton D., Warr G.W.;
 RT "Immunoglobulin heavy chain variable region gene evolution: structure
 and family relationships of two genes and a pseudogene in a teleost
 fish.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1566-1570(1988).
 DR PIR: B28966; B28966.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region; Signal.
 FT SIGNAL 19
 FT CHAIN 20 116 IG HEAVY CHAIN V REGION 5A.
 FT DOMAIN 20 49 FRAMEWORK 1.
 FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING 1.
 FT DOMAIN 55 68 FRAMEWORK 2.
 FT DOMAIN 69 84 COMPLEMENTARITY-DETERMINING 2.
 FT DOMAIN 85 116 FRAMEWORK 3.
 FT DTSUFID 41 114 BY SIMILARITY.
 FT NON_TER 116
 SQ SEQUENCE 116 AA; 12808 MW; 9C2279E2DF199B12 CRC64;

Query Match
 Best Local Similarity 95.8%; Score 46; DB 1; Length 116;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 20 EVOLVESGGL 30

RESULT 5
 HV3T_HUMAN STANDARD; PRT; 116 AA.
 AC P01781;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION GAL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75059123; Pubmed=4803843;
 RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
 RT "The primary structure of a monoclonal IgM-immunoglobulin
 (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-
 type), subgroup H III. Architecture of the complete IgM-molecule.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
 RN [2]
 RP REVISION TO THE COMPOSITION OF 28-33.
 RA Hilschmann N.;
 RL Submitted (JUN-1975) to the PIR data bank.
 CC -I- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S
 MACROGLOBULIN.
 DR PIR: A02064; M3HUGL.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; -.
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON_TER 116
 SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9AANA1282 CRC64;

Query Match
 Best Local Similarity 95.8%; Score 46; DB 1; Length 116;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 1 EVOLVESGGL 11

DB 1 EVOLVESGDL 11

RESULT 6
HV55_MOUSE STANDARD; PRT: 117 AA.
AC P18526:
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION 345 PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/CJ;
RX MEDLINE=89279149; PubMed=2499654;
RA Levy N.S., Malplero U.V., Lebecque S.G., Gearhart P.J.;
RT "Early onset of somatic mutation in Immunoglobulin VH genes during
the primary immune response."
RL J. Exp. Med. 169:2007-2019(1989).
CC -1- MISCELLANEOUS: THIS SEQUENCE BELONGS TO THE VH7183 SUBFAMILY.
DR PIR: J70502; HVMS34.
DR InterPro: IPR003006; -.
KW Pfam: PF00047; 19; 1.
FT SIGNAL 1 19 IG HEAVY CHAIN V REGION 345.
FT CHAIN 20 117 FRAMEWORK 1.
FT DOMAIN 20 49 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 50 54 FRAMEWORK 2.
FT DOMAIN 55 68 FRAMEWORK 3.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 86 117 FRAMEWORK 3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12902 MW; 49380E4627AC9A CRC64;

Query Match 95.8%; Score 46; DB 1; Length 117;
Best Local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
DB 20 EVOLVESGGL 30

RESULT 7
HV3V_HUMAN STANDARD; PRT: 118 AA.
ID HV3V_HUMAN
AC P80419;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION GAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=95255298; PubMed=7737190;
RA Stoppani M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-flavin monooxigenase
immunoglobulins."
RL Eur. J. Biochem. 228:886-893(1995).
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 118 118

SQ SEQUENCE 118 AA; 13087 MW; 6C21D810ED1B6D1F CRC64;

Query Match 95.8%; Score 46; DB 1; Length 118;
Best Local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
DB 1 EVOLVESGGL 11

RESULT 8
HV3P_HUMAN STANDARD; PRT: 119 AA.
ID HV3P_HUMAN
AC P01777;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION TEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=74142702; PubMed=4522793;
RA Capra J.D., Kehoe J.M.;
RT "Variable region sequences of five human immunoglobulin heavy chains
of the VH3 subgroup: definitive identification of four heavy chain
hypervariable regions."
RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM1 MYELOMA
PROTEIN.
DR PIR: A02060; G1HUTE.
DR HSSP: P01772; 2IG2.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 119 119
SQ SEQUENCE 119 AA; 12802 MW; 7E24DC852C7290A9 CRC64;

Query Match 95.8%; Score 46; DB 1; Length 119;
Best Local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
DB 1 EVOLVESGGL 11

RESULT 9
HV3E_HUMAN STANDARD; PRT: 120 AA.
ID HV3E_HUMAN
AC P01766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-III REGION BRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77117674; PubMed=65324;
RA Capra J.D., Hopper J.E.;
RT "Comparative studies on monocytic IgM lambda and IgG kappa from an
individual patient. III. The complete amino acid sequence of the VH
region of the IGM paraprotein."
RL Immunochemistry 13:995-999(1976).
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE

CC SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM
TYPE.

DR PIR: A02049; M3HUB.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13227 MW; D3F0428F7C2E6410 CRC64;

Query Match

Best Local Similarity 95.8%; Score 46; DB 1; Length 120;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 10

HV3D_HUMAN
ID HV3D_HUMAN STANDARD; PRT; 120 AA.
AC P01782;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-II REGION DOB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
RN [1]
RX MEDLINE:80020921; PubMed-114209;
RA Steiner L.A., Garcia Pardo A., Margolies M.N.;
RT "Amino acid sequence of the heavy-chain variable region of the
RT crystallizable human myeloma protein Dob.";
RL Biochemistry 18:4068-4080(1979).
RN [2]
RP CRYSTALLIZATION.
RA MEDLINE:80020920; PubMed-114208;
RA Steiner L.A., Lopes A.D.;
RT "The crystallizable human myeloma protein Dob has a hinge-region
RT deletion.";
RL Biochemistry 18:4054-4067(1979).
CC -1 MISCELLANEOUS: THIS GAMMA-1 MYELOMA PROTEIN HAS A DELETION IN THE
CC HINGE REGION. THERE ARE NO LIGHT-HEAVY OR INTER-HEAVY CHAIN
CC DISULFIDE BONDS.
DR PIR: A02065; GIHDB.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;

Query Match 95.8%; Score 46; DB 1; Length 120;
Best Local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 11

HV03_CAICR
ID HV03_CAICR STANDARD; PRT; 117 AA.
AC P03982;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V REGION G4 PRECURSOR.

GN G4.
OS Calman crocodilus (Spectacled calman) (Calman sclerops).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatoridae; Calman.
RN [1]
NCBI_TaxID=8499;
RP SEQUENCE FROM N.A.
RX MEDLINE:85140192; PubMed-2983316;
RA Litman G.W., Murphy K., Berger L., Litman R., Hinds K.,
RA Erickson B.W.;
RT "Complete nucleotide sequences of three VH genes in Calman, a
RT phylogenetically ancient reptile: evolutionary diversification in
RT coding segments and variation in the structure and organization of
RT recombination elements.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:844-848(1985).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: M12770; AAA9194.1; -
DR PIR: A02085; HVC0G4.
DR InterPro: IPR003006; -
DR Pfam: PF00047; 19; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117
FT DOMAIN 20 49
FT DOMAIN 50 54
FT DOMAIN 55 68
FT DOMAIN 69 85
FT DOMAIN 86 117
FT DISULFID 41 115
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12994 MW; 6330D7469AA55FC CRC64;

Query Match 91.7%; Score 44; DB 1; Length 117;
Best Local Similarity 72.7%; Pred. No. 0.031;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 20 EVOLVESGGL 30

RESULT 12
HV3D_HUMAN
ID HV3D_HUMAN STANDARD; PRT; 115 AA.
AC P01765;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG HEAVY CHAIN V-II REGION TIL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
RN [1]
NCBI_TaxID=9606;
RP SEQUENCE.
RX MEDLINE:78005528; PubMed-409716;
RA Wang A.-C., Wang I.Y., Fudenberg H.H.;
RT "Immunoglobulin structure and genetics. Identity between variable
RT regions of a mu and a gamma2 chain.";
RL J. Biol. Chem. 252:7193-7199(1977).
CC -1 MISCELLANEOUS: THE SEQUENCES OF THE V REGIONS OF THE HEAVY CHAINS
CC OF IGM AND IGG2 ISOLATED FROM A SINGLE PATIENT WITH BICLONAL
CC GAMMOPATHY ARE IDENTICAL. THEIR LIGHT CHAINS ARE APPARENTLY ALSO
CC IDENTICAL.

DR PIR: A02048; H3HUTL.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON-TER 115 115
 SO SEQUENCE 115 AA; 12356 MW; 4DCC67D179F62326 CRC64;

Query Match
 Best Local Similarity 89.6%; Score 43; DB 1; Length 115;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 1 EVOLVESGGL 11

RESULT 13
 HV3F_HUMAN STANDARD; PRT; 115 AA.
 AC P01767;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION BUT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=78137069; PubMed=416441;
 RA Torano A., Putnam F.W.;
 RT "Complete amino acid sequence of the alpha 2 heavy chain of a human
 RT 19A2 immunoglobulin of the A2m(2) allotype."
 RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
 CC -I- MISCELLANEOUS: THE SEQUENCE OF THE ALPHA-2, A2M(2) ALLOTYPE, C
 CC REGION OF THIS MYELOMA PROTEIN IS ALSO GIVEN.
 DR PIR: A02050; A2HUBU.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON-TER 115 115
 SO SEQUENCE 115 AA; 12379 MW; 208876A7DF52DCFA CRC64;

Query Match
 Best Local Similarity 89.6%; Score 43; DB 1; Length 115;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 1 EVOLVETGCGL 11

RESULT 14
 HV3S_HUMAN STANDARD; PRT; 115 AA.
 AC P01780;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION JUN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=74142702; PubMed=4522793;
 RA Capra J.D., Kehoe J.M.;
 RT "Variable region sequences of five human immunoglobulin heavy chains
 RT of the VH3 subgroup: definitive identification of four heavy chain

RT hypervariable regions."
 RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG3 MYELOMA
 CC PROTEIN.
 DR PIR: A02063; G3H0JN.
 DR HSSP: P01772; 21G2.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON-TER 115 115
 SO SEQUENCE 115 AA; 12563 MW; 68E668E531C12514 CRC64;

Query Match
 Best Local Similarity 89.6%; Score 43; DB 1; Length 115;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 1 DVOLVESGGL 11

RESULT 15
 HV3R_HUMAN STANDARD; PRT; 116 AA.
 AC P01779;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG HEAVY CHAIN V-III REGION TUR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=74142702; PubMed=4522793;
 RA Capra J.D., Kehoe J.M.;
 RT "Variable region sequences of five human immunoglobulin heavy chains
 RT of the VH3 subgroup: definitive identification of four heavy chain
 RT hypervariable regions."
 RL Proc. Natl. Acad. Sci. U.S.A. 71:845-848(1974).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGA1 MYELOMA
 CC PROTEIN.
 DR PIR: A02062; A1H0TU.
 DR InterPro: IPR003006; -
 DR Pfam: PF00047; 1g; 1.
 KW Immunoglobulin V region.
 FT NON-TER 116 116
 SO SEQUENCE 116 AA; 12431 MW; EB705F53A963F0C CRC64;

Query Match
 Best Local Similarity 89.6%; Score 43; DB 1; Length 116;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
 DB 1 EVOLVESGGL 11

Search completed: June 13, 2001, 14:30:37
 Job time: 527 sec

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035023; AAD56259.1;
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER
SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;

Query Match 95.8%; Score 46; DB 4; Length 118;
Best Local Similarity 90.9%; Pred. No. 0.094;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 3
Q9UL72 PRELIMINARY; PRT; 118 AA.
ID Q9UL72
AC Q9UL72;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035042; AAD56278.1;
DR HSSP; P01772; 2F84.
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER
SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

Query Match 95.8%; Score 46; DB 4; Length 118;
Best Local Similarity 90.9%; Pred. No. 0.094;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 4
Q9UL88 PRELIMINARY; PRT; 131 AA.
ID Q9UL88
AC Q9UL88;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035026; AAD56262.1;
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER
SQ SEQUENCE 131 AA; 14142 MW; 96E7D668E375DEA0 CRC64;

Query Match 95.8%; Score 46; DB 4; Length 131;
Best Local Similarity 90.9%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 5
Q9UL90 PRELIMINARY; PRT; 113 AA.
ID Q9UL90
AC Q9UL90;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035024; AAD56260.1;
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 19; 1.
FT NON_TER
SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;

Query Match 89.6%; Score 43; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESG 9
Db 1 EVOLVESG 9

RESULT 6
Q9UL71 PRELIMINARY; PRT; 121 AA.
ID Q9UL71
AC Q9UL71;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; AAD56279.1; -.
DR HSSP; P01772; 2F84.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; 1g; 1.
FT NON_TER 1 1
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13154 MW; 2F045CFA5D50736 CRC64;

Query Match 89.6%; Score 43; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGC 9
Db 1 EVOLVESGC 9

RESULT 7
O9UL84 ID O9UL84 PRELIMINARY; PRT; 122 AA.
AC O9UL84;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; AAD56279.1; -.
DR HSSP; P01772; 2F84.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; 1g; 1.
FT NON_TER 1 1
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 13579 MW; 36054D4136654588 CRC64;

Query Match 89.6%; Score 43; DB 4; Length 122;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGC 9
Db 1 EVOLVESGC 9

RESULT 8
O9UL93 ID O9UL93 PRELIMINARY; PRT; 116 AA.
AC O9UL93;
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; 1g; 1.
FT NON_TER 1 1
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

Query Match 79.2%; Score 38; DB 4; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VOLVESGC 9
Db 1 VOLVESGC 8

RESULT 9
O9UC53 ID O9UC53 PRELIMINARY; PRT; 16 AA.
AC O9UC53;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE 77 KDA SPONTANEOUS RECURRENT ABORTION-ASSOCIATED HUMAN EMBRYONIC
DE ANTIGEN/IGWHIT HOMOLOG (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=96033130; PubMed=8582963;
RA Shiraiishi Y., Shiraiishi Y., Yamamoto D., Hasegawa T., Kitamura W.,
RA Miki S., Tanaka T., Suzuki T., Soma H.;
RT "Diagnostic relevance of abortion-associated human embryonic antigen
RT expressed on the cell surface of tumour promoter-treated Bloom
RT syndrome cells.";
RL Hum. Reprod. 10:1694-1701(1995).
SQ SEQUENCE 16 AA; 1626 MW; C9C5ED2512FF3FB9 CRC64;

Query Match 77.1%; Score 37; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVOLVESGC 8
Db 1 EVOLVESGC 8

RESULT 10
O9UL94 ID O9UL94 PRELIMINARY; PRT; 119 AA.
AC O9UL94;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE HOMO-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035020; AAD56256.1; -
 DR INTERPRO: IPR003006; -
 DR PIRAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 13205 MW; 13E6AF5345F4A16E CRC64;

Query Match 77.1%; Score 37; DB 4; Length 119;
 Best Local Similarity 100.0%; Pred. No. 6.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVOLVESG 8
 Db 1 EVOLVESG 8

RESULT 11
 ID Q9UL92 PRELIMINARY; PRT; 124 AA.
 AC Q9UL92;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035022; AAD56258.1; -
 DR INTERPRO: IPR003006; -
 DR PIRAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 124 124
 SQ SEQUENCE 124 AA; 13580 MW; 1BAACBD96ACD2A2 CRC64;

Query Match 77.1%; Score 37; DB 4; Length 124;
 Best Local Similarity 100.0%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVOLVESG 8
 Db 1 EVOLVESG 8

RESULT 12
 ID Q9N0M4 PRELIMINARY; PRT; 124 AA.
 AC Q9N0M4;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ANTI-HUMAN A33 HEAVY CHAIN DOMAIN (FRAGMENT).
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Rader C., Ritzer G., Nathan S., Elia M., Gout I., Jungbluth A.A.,
 RA Cohen L.S., Melt S., Old L.J., Bardas C.F. III.;
 RT "The rabbit antibody repertoire as a novel source for the generation
 of therapeutic human antibodies.";
 RL J. Biol. Chem. 275:13668-13676(2000).
 DR EMBL: AF245503; AAF68450.1; -
 DR NON_TER 1 1
 DR NON_TER 124 124
 SQ SEQUENCE 124 AA; 13476 MW; 96D2B29FE27C24C8 CRC64;

Query Match 77.1%; Score 37; DB 6; Length 124;
 Best Local Similarity 88.9%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 3 OLVESGXL 11
 Db 3 OLVESGGL 11

RESULT 13
 ID Q9UL95 PRELIMINARY; PRT; 125 AA.
 AC Q9UL95;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035019; AAD56255.1; -
 DR INTERPRO: IPR003006; -
 DR PIRAM: PF00047; 19; 1.
 FT NON_TER 1 1
 FT NON_TER 125 125
 SQ SEQUENCE 125 AA; 13516 MW; 0D3CD5C232488EAC CRC64;

Query Match 77.1%; Score 37; DB 4; Length 125;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EVOLVESG 8
 Db 1 EVOLVESG 8

RESULT 14
 ID Q9VTG0 PRELIMINARY; PRT; 628 AA.
 AC Q9VTG0;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)

DE CG7628 PROTEIN.
 CN CG7628.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX MEDLINE=20196006; PubMed=10731132;
 RC STRAIN=BERKELEY;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Baller R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunker B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Modary C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Slieden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003546; AAF50089.1; -;
 DR FLYBASE: FBgn0036137; CG7628.
 DR INTERPRO: IPR001204; -;
 DR PFAM: PF01384; PHO4; 1.
 SQ SEQUENCE 628 AA; 66903 MW; 0A6D849EFDDBE1 CRC64;

Query Match 75.0%; Score 36; DB 5; Length 628;
 Best Local Similarity 70.0%; Pred. No. 67;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VOLVESGXL 11
 DB 372 LOVESGSL 381

RESULT 15
 ID 09ZXD1 PRELIMINARY; PRT; 92 AA.
 AC 09ZXD1;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE DNA, COMPLETE SEQUENCE.

OS Bacteriophage phi-105.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
 OC Lambda phage group.
 NCBI_TaxID=10717;
 RX MEDLINE=10717;
 RC [1]
 RA Kobaishi K., Okamura K., Inoue T., Sato T., Kobayashi Y.;
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB016282; BAA36661.1; -;
 DR INTERPRO: IPR001387; -;
 DR PFAM: PF01381; HTH_3; 1.
 SQ SEQUENCE 92 AA; 10362 MW; 2BDD192A01BE2FBD CRC64;

Query Match 72.9%; Score 35; DB 9; Length 92;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESG 9
 DB 71 EIQVEEG 79

Search completed: June 13, 2001, 14:29:45
 Job time: 546 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:27:07 ; Search time 78.71 Seconds

(without alignments)
2.685 Million cell updates/sec

Title: PCT-US01-05825A-28

Perfect score: 48

Sequence: 1 EVOLVESGXL 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	97.9	119	3	US-08-545-809A-102
2	47	97.9	150	2	US-08-345-321-6
3	46	95.8	11	2	US-08-859-931A-7
4	46	95.8	18	1	US-08-331-398A-62
5	46	95.8	18	2	US-08-331-397B-62
6	46	95.8	18	2	US-08-739-804A-61
7	46	95.8	20	1	US-08-050-113-1
8	46	95.8	20	2	US-08-859-931A-2
9	46	95.8	26	1	US-08-471-780C-80
10	46	95.8	26	1	US-08-467-282B-80
11	46	95.8	26	2	US-08-471-282A-80
12	46	95.8	26	2	US-08-466-710C-80
13	46	95.8	26	3	US-08-468-739C-80
14	46	95.8	30	1	US-07-977-696C-75
15	46	95.8	30	1	US-08-129-930B-75
16	46	95.8	30	2	US-08-470-139-17
17	46	95.8	35	2	US-08-765-179B-1
18	46	95.8	98	1	US-08-211-202-116
19	46	95.8	98	1	US-07-942-245-37
20	46	95.8	98	2	US-08-665-202-31
21	46	95.8	100	1	US-08-320-515B-2
22	46	95.8	100	1	US-08-320-515B-3
23	46	95.8	100	1	US-08-309-025-2
24	46	95.8	100	1	US-08-309-025-3
25	46	95.8	108	1	US-08-428-197-4
26	46	95.8	108	5	PCT-US93-10555-4
27	46	95.8	109	2	US-08-379-057-32

28	46	95.8	109	2	US-08-428-197-3	Sequence 3, Appl1
29	46	95.8	109	5	PCT-US93-10555-3	Sequence 3, Appl1
30	46	95.8	111	3	US-08-545-809A-121	Sequence 121, App
31	46	95.8	113	3	US-08-974-899-6	Sequence 6, Appl
32	46	95.8	114	2	US-08-887-352B-11	Sequence 11, Appl
33	46	95.8	114	2	US-08-887-352B-12	Sequence 12, Appl
34	46	95.8	114	3	US-08-545-809A-124	Sequence 124, App
35	46	95.8	114	4	US-09-109-207C-11	Sequence 11, Appl
36	46	95.8	114	4	US-09-109-207C-12	Sequence 11, Appl
37	46	95.8	115	2	US-08-379-057-31	Sequence 31, Appl
38	46	95.8	115	2	US-08-545-809A-122	Sequence 122, App
39	46	95.8	115	4	US-08-767-128-36	Sequence 36, Appl
40	46	95.8	116	1	US-08-478-039-74	Sequence 74, Appl
41	46	95.8	116	1	US-08-478-039-103	Sequence 103, Appl
42	46	95.8	116	1	US-08-476-349A-74	Sequence 74, Appl
43	46	95.8	116	1	US-08-476-349A-103	Sequence 103, App
44	46	95.8	116	2	US-08-428-197-10	Sequence 10, Appl
45	46	95.8	116	5	PCT-US93-10555-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-545-809A-102
Sequence 102, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545, 809A
FILING DATE: 27-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29, 066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-102

Query Match 97.9%; Score 47; DB 3; Length 119;
Best Local Similarity 90.9%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVOLVESGXL 11

Db 20 EVOLVESGAL 30

RESULT 2

US-08-345-321-6

Sequence 6, Application US/08345321

Patent No. 5914109

GENERAL INFORMATION:

APPLICANT: ZOLLA-PAZNER, Susan

TITLE OF INVENTION: HETEROBIRIDOMAS PRODUCING HUMAN

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Browdy and Neimark

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/345,321

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/872,675

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Browdy, Roger L.

REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: ZOLLA-PAZNER1B

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

TELEX: 248633

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 150 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-345-321-6

Query Match 97.9% Score 47; DB 2; Length 150;
Best Local Similarity 90.98; Pred. No. 0.063;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11

Db 20 EVOLVESGAL 30

RESULT 3

US-08-859-931A-7

Sequence 7, Application US/08859931A

Patent No. 5945510

GENERAL INFORMATION:

APPLICANT: PASANO, Alessio

TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A

TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W., Suite 800

CITY: Washington, D.C.

CITY: Washington, D.C.

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/859,931A

FILING DATE: 21 MAY 1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: KIT, Gordon

REGISTRATION NUMBER: 30,764

REFERENCE/DOCKET NUMBER: A-6901

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 293-7860

TELEFAX: (202) 293-7860

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

US-08-859-931A-7

Query Match 95.8% Score 46; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0062;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11

Db 1 EVOLVESGXL 11

QY 1 EVOLVESGXL 11


```

: ATTORNEY/AGENT INFORMATION:
: NAME: Hunter, Tom
: REGISTRATION NUMBER: 38,438
: REFERENCE/DOCKET NUMBER: 015280-12611005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 543-9600
: TELEFAX: (415) 543-5043
: INFORMATION FOR SEQ ID NO: 62:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 18 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: OS-08-331-998A-62

```

Query Match	95.8%	Score 46:	DB 1:	Length 18:
Best Local Similarity	90.9%	Pred. No.	0.01:	
Matches 10:	Conservative	0:	Mismatches	1: Indels 0: Gaps 0:

RESULT 5
 US-08-331-397B-62
 : Sequence 62, Application US/08331397B
 Patent No. 5981726
 GENERAL INFORMATION:
 APPLICANT: Pastan, Ira
 APPLICANT: Benhar, Itai
 TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-
 TITLE OF INVENTION: Specific Antibody Fragments, Fusion Proteins, and Uses
 TITLE OF INVENTION: thereof
 NUMBER OF SEQUENCES: 68
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Crew
 STREET: One Market Plaza, Stewart Street Plaza
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105-1492
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/331,397B
 FILING DATE: 28-OCT-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/767,331
 FILING DATE: 30-SEP-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/596,289
 FILING DATE: 12-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Hunter, Tom
 REGISTRATION NUMBER: 38,498
 REFERENCE/DOCKET NUMBER: 015280-126120US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 543-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 62:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: Linear
 MOLECULE TYPE: peptide

US-08-331-397B-62

Query Match	95.8%	Score 46:	DB 2:	Length 18:
Best Local Similarity	90.9%	Pred. No. 0.01:		
Matches 10:	Conservative	0:	Mismatches 1:	Indels 0:
Qy	1	EVOLVESGXL 11		
Db	1	EVOLVESGGGL 11		

RESULT 6
 US-08-759-804A-61
 Sequence 61 Application US/08759804A
 Patent No. 5990296
 GENERAL INFORMATION:
 APPLICANT: Pastan, Ira
 APPLICANT: Willingham, Mark
 APPLICANT: Fitzgerald, David J.
 APPLICANT: Brinkmann, Ulrich
 APPLICANT: Pal, Lee
 TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
 TITLE OF INVENTION: Fusion Proteins, and Uses Thereof
 NUMBER OF SEQUENCES: 68
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/759,804A
 FILING DATE: 03-DEC-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/331,398
 FILING DATE: 28-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/767,331
 FILING DATE: 30-SEP-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/596,289
 FILING DATE: 12-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Weber, Ellen L.
 REGISTRATION NUMBER: 32,762
 REFERENCE/DOCKET NUMBER: 015280-126140US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 61:
 SEQUENCE CHARACTERISTICS
 LENGTH: 18 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-759-804A-61

```

Query Match          95.8%;   Score 46;   DB 2;   Length 18;
Best Local Similarity 90.9%;   Pred. NO. 0.01;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

Oy      1 EVOLVSGGXL 11
||||||| |

```

Db 1 EVOLVESGGL 11

RESULT 7

US-08-050-113-1

Sequence 1, Application US/08050113

Patent No. 551454

GENERAL INFORMATION:

APPLICANT: Bacus, Sarah S

APPLICANT: Yarden, Yosef

APPLICANT: Sela, Michael

TITLE OF INVENTION: Methods and Compositions for Cancer

TITLE OF INVENTION: Therapy and for Prognosticating Responses to Cancer

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Richard J. Rodrick, Becton, Dickinson &

STREET: 1 Becton Drive

CITY: Franklin Lakes

STATE: New Jersey

COUNTRY: USA

ZIP: 07417-1880

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/050,113

FILING DATE: 07-OCT-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Highest, David W

REGISTRATION NUMBER: 30,265

REFERENCE/DOCKET NUMBER: P-2624

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 847 5317

TELEFAX: 201 848 9228

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-08-050-113-1

Query Match 95.8%; Score 46; DB 1; Length 20;

Best Local Similarity 90.9%; Pred. No. 0.012;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGGL 11

Db 1 EVOLVESGGL 11

RESULT 8

US-08-859-931A-2

Sequence 2, Application US/08859931A

Patent No. 5945510

GENERAL INFORMATION:

APPLICANT: FASANO, Alessio

TITLE OF INVENTION: SUBSTANTIALLY PURE ZONULIN, A

TITLE OF INVENTION: PHYSIOLOGICAL MODULATOR OF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W., Suite 800

CITY: Washington, D.C.

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/859,931A

FILING DATE: 21 MAY 1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: KIT, Gordon

REGISTRATION NUMBER: 30,764

REFERENCE/DOCKET NUMBER: A-6901

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 293-7060

TELEFAX: (202) 293-7860

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHEICAL: NO

US-08-859-931A-2

Query Match 95.8%; Score 46; DB 2; Length 20;

Best Local Similarity 90.9%; Pred. No. 0.012;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGGL 11

Db 1 EVOLVESGGL 11

RESULT 9

US-08-471-780C-80

Sequence 80, Application US/08471780C

Patent No. 5759808

GENERAL INFORMATION:

APPLICANT: Castelman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,780C

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106,944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potler, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-780C-80

Query Match 95.8%; Score 46; DB 1; Length 26;
Best Local Similarity 90.9%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 10
US-08-467-282B-80
Sequence 80, Application US/08467282B
Patent No. 5800988
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,282B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
US-08-471-780C-80

ORGANISM: Camelus dromedarius
US-08-467-282B-80

Query Match 95.8%; Score 46; DB 1; Length 26;
Best Local Similarity 90.9%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 11
US-08-471-282A-80
Sequence 80, Application US/08471282A
Patent No. 5840853
GENERAL INFORMATION:
APPLICANT: Casterman, Cecile
TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
NUMBER OF SEQUENCES: 130
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,282A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/106,944
FILING DATE: 17-AUG-1993
APPLICATION NUMBER: FR 92402326.0
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93401310.3
FILING DATE: 21-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E.R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 04958.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Camelus dromedarius
US-08-471-282A-80

Query Match 95.8%; Score 46; DB 2; Length 26;
Best Local Similarity 90.9%; Pred. No. 0.015;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVOLVESGXL 11
Db 1 EVOLVESGGL 11

RESULT 12

US-08-466-710C-80

Sequence 80, Application US/08466710C

Patent No. 5874541

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pinegan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,710C

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/106,944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958.0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 26 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Camelus dromedarius

US-08-466-710C-80

Query Match

Best Local Similarity 95.8%; Score 46; DB 2; Length 26;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11

DB 1 EVOLVESGGL 11

RESULT 13

US-08-468-739C-80

Sequence 80, Application US/08468739C

Patent No. 6015695

GENERAL INFORMATION:

APPLICANT: Casterman, Cecile

APPLICANT: Hamers, Raymond

TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains

NUMBER OF SEQUENCES: 130

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pinegan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,739C

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/106,944

FILING DATE: 17-AUG-1993

APPLICATION NUMBER: FR 92402326.0

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93401310.3

FILING DATE: 21-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E.R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 04958.0008-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 26 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Camelus dromedarius

US-08-468-739C-80

Query Match

Best Local Similarity 95.8%; Score 46; DB 3; Length 26;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11

DB 1 EVOLVESGGL 11

RESULT 14

US-07-977-696C-75

Sequence 75, Application US/07977696C

Patent No. 5792852

GENERAL INFORMATION:

APPLICANT: do Couto, Fernando J.R.

APPLICANT: Ceriani Dr., Roberto L.

APPLICANT: Peterson Dr., Jerry A.

APPLICANT: Padian Dr., Eduardo A.

TITLE OF INVENTION: Analogous Peptides with Specificity

TITLE OF INVENTION: for Carcinomas and Kit and Diagnostic Vaccination

NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI

STREET: 444 South Flower Street, Suite 2000

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977.696C
FILING DATE: 11-16-92
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Amzel Ph.D., Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38227
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 748-6868
TELEFAX: (510) 748-6688
TELEX: P.A.
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-977-696C-75

Query Match 95.8%; Score 46; DB 1; Length 30;
Best Local Similarity 90.9%; Pred. No. 0.018;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
|||||||
DB 1 EVOLVESGGL 11

RESULT 15
US-08-129-930B-75
Sequence 75, Application US/08129930B
Patent No. 5804187
GENERAL INFORMATION:
APPLICANT: do Couto Dr., Fernando J.R.
APPLICANT: Ceriani Dr., Roberto L.
APPLICANT: Peterson Dr., Jerry A.
APPLICANT: Padlan Dr., Eduardo A.
TITLE OF INVENTION: Analogue Peptides With Broad
TITLE OF INVENTION: Carcinoma Specificity, and Kit and
TITLE OF INVENTION: Diagnostic Vaccination and
TITLE OF INVENTION: Therapeutic Methods
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESS: V. AMZEL & ASSOC.
STREET: 2055 No. 5804187th Broadway, Suite 201
CITY: Walnut Creek
STATE: California
COUNTRY: USA
ZIP: 94596
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/129.930B
FILING DATE: September 30, 1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Amzel Ph.D., Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: CRFCC-008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 521-1333
TELEFAX: (510) 521-3541
TELEX: n.a.
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-129-930B-75

Query Match 95.8%; Score 46; DB 1; Length 30;
Best Local Similarity 90.9%; Pred. No. 0.018;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVOLVESGXL 11
|||||||
DB 1 EVOLVESGGL 11

Search completed: June 13, 2001, 14:27:07
Job time: 628 sec

Wed Jun 13 15:00:40 2001

pct-us01-05825a-28.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:47 ; Search time 150.28 Seconds

(without alignments)
3.043 Million cell updates/sec

Title: PCT-US01-05825A-29

Perfect score: 42

Sequence: 1 VGVLRPG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	8	21	Y84662
2	42	100.0	8	21	Y84662
3	37	88.1	4472	18	W22601
4	35	83.3	119	21	B40443
5	35	83.3	4545	18	W22611
6	35	83.3	4550	18	W23716
7	35	83.3	4550	18	W22606
8	34	81.0	499	21	G11597
9	34	81.0	499	21	G52088
10	34	81.0	564	21	G11596
11	34	81.0	564	21	G52087

12	34	81.0	616	21	G52086
13	34	81.0	617	21	G11595
14	34	81.0	785	18	W16303
15	33	78.6	467	20	V35025
16	33	78.6	807	19	W70461
17	33	78.6	807	19	W70463
18	33	78.6	807	19	W70465
19	33	78.6	1066	21	B07561
20	33	78.6	1218	21	B07563
21	33	78.6	1832	21	V58575
22	33	78.6	1838	15	R53257
23	33	78.6	4572	19	W52845
24	32	76.2	150	20	V76536
25	32	76.2	202	21	B58805
26	32	76.2	218	22	B52452
27	32	76.2	318	21	V83931
28	32	76.2	428	21	G21068
29	32	76.2	433	20	V18055
30	32	76.2	433	20	V18056
31	32	76.2	456	21	G21067
32	32	76.2	579	21	G53282
33	32	76.2	584	20	V17507
34	32	76.2	610	21	B14341
35	32	76.2	610	21	G53281
36	32	76.2	626	21	G53280
37	32	76.2	806	19	W53103
38	32	76.2	1511	21	B28182
39	31	73.8	53	21	B40807
40	31	73.8	68	20	V74207
41	31	73.8	194	21	B58439
42	31	73.8	247	20	V76636
43	31	73.8	247	20	V73971
44	31	73.8	248	21	V54422
45	31	73.8	362	21	B43555

ALIGNMENTS

RESULT 1	
ID	Y84662 standard; Protein; 8 AA.
XX	
XX	Y84662:
DT	25-JUL-2000 (first entry)
XX	
DE	Peptide antagonist FZ1/1 of zonula occludens toxin (zot) polypeptide.
KW	Human; zot; zonula occludens toxin; zonulin; antigen presenting cell;
KW	APC; lymphocyte proliferation; antigen; auto-immune disorder;
KW	immune-related disorder; immune system rejection; multiple sclerosis;
KW	organ transplantation; inflammatory disease; allergic disease;
KW	rheumatoid arthritis; insulin dependent diabetes mellitus;
KW	celiac disease; Sjogren's syndrome; systemic lupus erythematosus;
KW	auto-immune thyroiditis; idiopathic thrombocytopenic purpura;
KW	hemolytic anemia; Grave's disease; Addison disease; autoimmune orchitis;
KW	pernicious anemia; vasculitis; autoimmune coagulopathy; polymyositis;
KW	myasthenia gravis; polyneuritis; pemphigus; rheumatic carditis;
KW	dermatomyositis; scleroderma; asthma; psoriasis; eczematous dermatitis;
KW	Kaposi's sarcoma; inflammatory bowel disease; proliferative disorder;
KW	PCR primer: SS.
XX	
XX	Synthetic.
OS	
XX	
PN	WO200015252-A1.
XX	
PD	23-MAR-2000.
XX	
PF	09-SEP-1999; 99WO-US18842.
XX	
PR	14-SEP-1998; 98US-0100266.
XX	

Fasano A, Szteln MB, Lu R, Tanner MK, F1 XX

DR WP1; 2000-271257/23
XX

PT Suppression of antigen presenting cell mediated lymphocyte
PT proliferation, by administering a Zot-related immunoregulator useful
PT for treating immune-related disorders, immune system rejection
PT subsequent to tissue or organ transplantation -
XX

Example 1; Page 59; 95pp; English.
PS
XX

The present sequence represents a peptide antagonist of zonula occludens toxin (zot) polypeptide. The specification describes a method of suppressing antigen presenting cell (APC)-mediated lymphocyte proliferation in a mammalian host pre-exposed to a particular antigen. The method comprises administering to the host an effective amount of a zot-related immunoregulator selected from zot (zonula occludens toxin) or zonulin, the amount effective to down-regulate the activity of the APC. The method can be used to down-regulate APC-mediated lymphocyte proliferation in mammalian hosts suffering from auto-immune or immune-related disorders, immune system rejection subsequent to tissue or organ transplantation, or inflammatory or allergic diseases. The auto-immune or immune related disorders include multiple sclerosis, rheumatoid arthritis, insulin dependent diabetes mellitus, celiac thyroiditis, Sjogren's syndrome, systemic lupus erythematosus, auto-immune Grave's disease, Addison disease, auto-immune orchitis, peritonitis anemia, vasculitis, autoimmune coagulopathies, myasthenia gravis, polyneuritis, pemphigus, rheumatic carditis, polymyositis, dermatomyositis, and scleroderma. The inflammatory or allergic diseases include disorders selected from asthma, psoriasis, eczematous dermatitis, Kaposi's sarcoma, multiple sclerosis, inflammatory bowel disease, proliferative disorders of smooth muscle cells, and inflammatory conditions associated with mycotic, viral, parasitic, or bacterial infections.

Query Match	100.0%;	Score 42;	DB 21;	Length 8
Best Local Similarity	100.0%;	Pred. No. 3.2e+05;		
Matches	8;	Conservative 0.	Mismatches 0	

Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
--------------	----	------------	----	--------	----	------	---

vy	1	VGVLGRPG	8
Db	1	vgvlgrrpg	8

RESULT	2
Y79137	
ID	Y79137 standard; Peptide; 8 AA
XX	

05-JUN-2000 (first entry)

Peptide FZI/1.

zonulin; antagonist; zonula occludens toxin receptor
human; blood-brain barrier; antiinflammatory;
gastrointestinal inflammation; therapy.

Synthetic.

WO200007609-A1

17-FEB-2000

28-JUL-1999; 99MO-JS16683

03-AUG-1998. 0815-0133015

03-AUG-1998; 98US-0127815.

XX
XX
PA
XX

(UYMA-) UNIV MARYLAND BALTIMORE

PI Fasano A
XX

DR WPI; 2000-205565/18
XX

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 XX
 XX
 PS Example 4; Page 38; 69pp; English.
 XX

Example 4; page 38; 69pp; English

The present sequence is that of peptide F171, which is based on a peptide motif (see V79119) of human adult heart zonulin, but has 3 amino acid changes, corresponding to Gly at position 8, Val at position 12 and Gln at position 13 of the adult heart zonulin N-terminal sequence (see V79130). It was used in an experiment to demonstrate that there is a region spanning residues 8-15 of the N-terminal end of zonulin family proteins and the zonula occludens toxin (ZOT) that is crucial for binding to the target receptors that the residues in positions 8, 12 and 13 determine the tissue specificity of this binding. The N-terminal sequences of human adult and foetal zonulins (see V79130-36) were compared ZOT to identify the motif involved in receptor binding. Peptide antagonists (see V79105-29) based on this motif are useful as antiinflammatory agents for treatment of gastrointestinal inflammation, and for treatment of conditions associated with breakdown of the blood-brain barrier.

Sequence	8 AA
SSQ	

Query Match	100.0%;	Score 42;	DB 21;	Length 8
Best Local Similarity	100.0%;	Pred. No. 3.2e+05;		
Matches	8;	Conservative 0;	Mismatches 0;	

1 VGVLGKPG 8

8 vgl. Grp 1. I

RESULT	3
222601	
W22601	standard; Protein; 4472 AA
XX	

W22601;

27-FEB-1998 (first entry)

Tylactone synthase ORF1 protein

tylactone synthase; tylg gene; multifunctional protein, polyketide; tyllactone synthesis; antibiotic; tylosin.

Streptomyces fradiae

Key	Location/Qualifiers
Miscellaneous	

encoded by GTC

43:144 /
/note="ketosynthase domain kcs/v

586., 914

Domain	$\nu_{\text{max}}/\text{m}(\text{s})$
954..1038	
(note - "a")	

Domain	Accession	Protein domain
1060..1483		

Domain	Accession number
7-nucleotide consensus sequence	U00001
1606..1934	U00001

Domain	/note="acyltransferase domain, AT1 2199..2378
--------	--

Domain	/note="ketoreductase domain, KR1"
2469	3553


```

FT      Domain      /note="acyl carrier protein domain, ACP1"
FT      2576..2999
FT      /note="ketosynthase domain, KS2"
FT      3149..3477
FT      Domain      /note="acyltransferase domain, AT2"
FT      3502..3687
FT      Domain      /note="dehydratase domain, DH2"
FT      4016..4200
FT      /note="ketoreductase domain, KR2"
FT      4302..4385
FT      Domain      /note="acyl carrier protein domain, ACP2"
FT      4302..4385
XX      PF791655-A2.
XX      27-AUG-1997.
XX      19-FEB-1997; 97EP-0301056.
XX      22-FEB-1996; 96US-0012078.
XX      (BLIT) LILLY & CO ELI.
XX      Dehoff BS, Kuhstoss SA, Rosteck PR, Sutton KL;
XX      WPI: 1997-418046/39.
XX      DR      N-PSDB; T80413.
XX      PT      DNA encoding Streptomyces fradiae tyactone synthase domain - for
XX      production of tylosin-related polypeptide compounds
XX      Claim 9; Pages 66-80; 220pp; English.
XX      PS      W22601-W22605 represent proteins encoded by the tyactone synthase gene
XX      cluster of the invention. The gene cluster is also referred to as the
XX      tylG gene, and was isolated from Streptomyces fradiae. These sequences
XX      are multifunctional proteins which direct the synthesis of the polypeptide
XX      tyactone, isolated from Streptomyces fradiae. Tyactone is the basic
XX      building block of the antibiotic tylosin. The DNA sequence can be
XX      modified so as to alter the type of carboxylic acids incorporated, the
XX      number of carboxylic acids incorporated and/or the post-condensation
XX      reactions performed, thereby resulting in novel tylosin-related
XX      polypeptides.
XX      CC      CC
XX      SO      Sequence 4472 AA:
XX
XX      Query Match      88.1%; Score 37; DB 18; Length 4472;
XX      Best Local Similarity 85.7%; Pred. No. 6.6e+02;
XX      Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      2 GVUGRPG 8
XX      1:|||||
XX      DB      4155 g11grpg 4161
XX
XX      RESULT 4
XX      ID      B40443 standard; Protein; 119 AA.
XX      AC      B40443;
XX      DT      08-FEB-2001 (first entry)
XX
XX      Human ORFX ORF207 polypeptide sequence SEQ ID NO:414.
XX
XX      Human: open reading frame: ORFX; detection; cytosstatic; hepatotropic;
XX      vulnary; antiposrotic; antiparkinsonian; nootropic; neuroprotective;
XX      anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX      immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX      hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX      antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
XX      antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX      neurodegenerative disorder; osteoarthritis; graft vs host disease;

```

```

KW      cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW      cholesterol ester storage; systemic lupus erythematosus; infection;
KW      severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW      allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW      bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW      thrombosis; contraceptive.
XX
XX      OS      Homo sapiens.
XX      PN      W0200058473-A2.
XX      PD      05-OCT-2000.
XX
XX      PF      31-MAR-2000; 2000WO-US08621.
XX      PR      31-MAR-1999; 99US-0127607.
XX      PR      02-APR-1999; 99US-0127636.
XX      PR      05-APR-1999; 99US-0127728.
XX      PR      30-MAR-2000; 2000US-0540763.
XX
XX      PA      (CURA-) CURAGEN CORP.
XX      PI      Shinkets RA, Leach M;
XX      DR      WPI: 2000-602362/57.
XX      DR      N-PSDB; C74652.
XX      PT      Novel nucleic acids and peptides derived from open reading frame X,
XX      useful for treating e.g. cancers, proliferative disorders,
XX      neurodegenerative disorders and cardiovascular disease -
XX      Claim 11; Page 640; 5507pp; English.
XX      PS      C74446 to C77606 encode the proteins given in B40237 to B43397, which
XX      represent the human ORFX open reading frames 1 to 3161. The ORFX
XX      sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX      antiposrotic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX      anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX      cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX      dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX      antiviral; antifungal; antirheumatic; antihypoid; and antinaemic. The
XX      sequences can be used for determining the presence of or predisposition
XX      to, or preventing or treating pathological conditions associated with an
XX      ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX      proteins in gene therapy vectors. The proteins and nucleic acids may be
XX      used to treat cancers, proliferative disorders, neurodegenerative
XX      disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX      diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX      storage, systemic lupus erythematosus, severe combined immunodeficiency
XX      (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX      disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX      cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX      enhance coagulation; to inhibit thrombosis; and as a contraceptive.
XX      CC      CC
XX      SO      Sequence 119 AA:
XX
XX      Query Match      83.3%; Score 35; DB 21; Length 119;
XX      Best Local Similarity 87.5%; Pred. No. 48;
XX      Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX      QY      1 VGVLGPRG 8
XX      11111111
XX      DB      31 vgvlgppg 38
XX
XX      RESULT 5
XX      ID      W22611 standard; Protein; 4545 AA.
XX      AC      W22611;
XX      DT      02-MAR-1998 (first entry)

```

XX DE Hybrid srmg/tylg ORF1 protein.
 XX XX Ty lactone synthase gene cluster; tylg gene; multifunctional protein;
 KW polyketide; ty lactone synthesis; antibiotic; tylosin; hybrid protein.
 XX OS Streptomyces ambofaciens.
 OS Streptomyces fradiae.
 FH FH Key
 FT FT Modified-site
 FT FT 1 Location/Qualifiers
 FT FT 15..418 /note="encoded by GTC"
 FT FT Domain
 FT FT 525..882 /note="ketosynthase domain, KS'(s), from tylg"
 FT FT 942..1025 /note="acyltransferase domain, AT(s), from tylg"
 FT FT Domain
 FT FT 1060..1483 /note="acyl carrier protein domain, ACP(s), from tylg"
 FT FT 1596..1953 /note="ketosynthase domain, KS1, from srmg"
 FT FT Domain
 FT FT 2232..2416 /note="acyltransferase domain, AT1, from srmg"
 FT FT 2533..2616 /note="ketoreductase domain, KR1, from srmg"
 FT FT Domain
 FT FT 2641..3064 /note="acyl carrier protein domain, ACP1, from srmg"
 FT FT 3184..3520 /note="ketosynthase domain, KS2, from srmg"
 FT FT Domain
 FT FT 3546..3727 /note="acyltransferase domain, AT2, from srmg"
 FT FT 4083..4268 /note="dehydratase domain, DH2, from srmg"
 FT FT Domain
 FT FT 4374..4457 /note="ketoreductase domain, KR2, from srmg"
 FT FT Domain
 FT FT /note="acyl carrier protein domain, ACP2, from srmg"
 XX XX EP791655-A2.
 PN 27-AUG-1997.
 PD 19-FEB-1997; 97EP-0301056.
 XX 22-FEB-1996; 96US-0012078.
 PR (ELIL) LILLY & CO ELI.
 PA Dehoff BS, Kunstoss SA, Rostock PR, Sutton KL;
 XX N-PSDB; T80415.
 DR WPI; 1997-418046/39.
 XX DNA encoding Streptomyces fradiae ty lactone synthase domain - for
 PT production of tylosin-related polyketide compounds
 XX Claim 23; Pages 198-212; 220pp; English.
 PS This sequence represents a hybrid protein of the invention. This sequence
 CC was created by replacing a EcoRI-ApaI fragment of srmg ORF1 with a
 CC EcoRI-SpuI fragment from tylg ORF1. The position of the residues from
 CC each of the two genes is not given in the specification. The srmg gene
 CC (see T80414) was isolated from Streptomyces ambofaciens, and encodes the
 CC multi-functional proteins which direct the synthesis of the polyketide
 CC platenolide. Platenolide is the basic building block of the macroide
 CC antibiotic spiramycin. The tylg gene (see T80413) is the ty lactone
 CC synthase gene cluster of the invention. The tylg sequence was isolated
 CC from Streptomyces fradiae, and encodes multifunctional proteins which
 CC direct the synthesis of the polyketide ty lactone. Ty lactone is the basic
 CC building block of the antibiotic tylosin. The hybrid sequence can be used
 CC to transform S. ambofaciens lacking the srmg ORF1 sequence, or S. fradiae
 CC lacking the tylg ORF1 sequence, so that they can produce polyketides. The
 CC DNA sequence can be modified so as to alter the type of carboxylic acids
 CC incorporated, the number of carboxylic acids incorporated and/or the

CC post-condensation reactions performed, thereby resulting in novel
 CC tylosin-related polyketides.
 XX SQ Sequence 4545 AA;
 Query Match 83.3%; Score 35; DB 18; Length 4545;
 Best Local Similarity 85.7%; Pred No. 1.3e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GVLGRP 8
 Db 4218 g11grp9 4224
 RESULT 6
 W23716
 ID W23716 standard; Protein; 4550 AA.
 XX W23716;
 AC 27-FEB-1998 (first entry)
 DT 27-FEB-1998 (first entry)
 DE Platenolide synthase ORF1 protein.
 XX Platenolide synthase gene cluster; platenolide production; srmg gene;
 KW multi-functional protein; macroide antibiotic; spiramycin.
 XX OS Streptomyces ambofaciens.
 OS Streptomyces ambofaciens.
 FH FH Key
 FT FT Modified-site
 FT FT 1 Location/Qualifiers
 FT FT 15..418 /note="encoded by GTC"
 FT FT Domain
 FT FT 525..882 /note="ketosynthase domain, KS'(s)"
 FT FT 942..1025 /note="acyltransferase domain, AT(s)"
 FT FT Domain
 FT FT 1060..1483 /note="acyl carrier protein domain, ACP(s)"
 FT FT 1596..1953 /note="ketosynthase domain, KS1"
 FT FT Domain
 FT FT 2232..2416 /note="acyltransferase domain, AT1"
 FT FT 2533..2616 /note="ketoreductase domain, KR1"
 FT FT Domain
 FT FT 2641..3064 /note="acyl carrier protein domain, ACP1"
 FT FT 3184..3520 /note="ketosynthase domain, KS2"
 FT FT Domain
 FT FT 3546..3727 /note="acyltransferase domain, AT2"
 FT FT 4083..4268 /note="dehydratase domain, DH2"
 FT FT Domain
 FT FT 4374..4457 /note="ketoreductase domain, KR2"
 FT FT Domain
 FT FT /note="acyl carrier protein domain, ACP2"
 XX EP791656-A2.
 PN 27-AUG-1997.
 PD 19-FEB-1997; 97EP-0301066.
 XX 22-FEB-1996; 96US-0012050.
 PR (ELIL) LILLY & CO ELI.
 PA Burgett SG, Kunstoss SA, Rao RN, Richardson MA;
 XX Rostock PR;
 XX WPI; 1997-418047/39.
 DR N-PSDB; T78508.

XX DNA encoding Streptomyces ambofaciens platenolide synthase domain -
 PT for production of spiramycin-related polyketide antibiotics
 PS Claim 8; Pages 33-47; 81pp; English.
 CC W23716-W23720 represent proteins encoded by the platenolide synthase gene
 CC cluster of the invention. The gene cluster is also referred to as the
 CC srmg gene, and was isolated from Streptomyces ambofaciens. These
 CC sequences are multi-functional proteins which direct the synthesis of the
 CC polyketide platenolide. Platenolide is the basic building block of the
 CC macroide antibiotic spiramycin. The DNA can be used to produce compounds
 CC exhibiting antibiotic activity based on the platenolide structure,
 CC including specifically the macroide antibiotic spiramycin and spiramycin
 CC analogues and derivatives. Modifications of the platenolide synthase DNA
 CC sequence can be made so as to change the number and type of carboxylic
 CC acids incorporated into the growing polyketide chain and to change the
 CC kind of post-condensation processing that is conducted.
 CC XX
 SO Sequence 4550 AA;

Query Match 83.3%; Score 35; DB 18; Length 4550;
 Best Local Similarity 85.7%; Pred. No. 1.5e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLGRPG 8
 Db 4223 gllgrpg 4229
 ||:|||||

RESULT 7
 W22606 W22606 standard; Protein; 4550 AA.
 XX
 AC W22606;
 XX
 DT 27-FEB-1998 (first entry)
 DE Platenolide synthase ORF1 protein.
 XX
 KW Ty lactone synthase gene cluster; tylg gene; multifunctional protein;
 KM polyketide; ty lactone synthesis; antibiotic; tylosin.
 XX
 OS Streptomyces ambofaciens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note- "encoded by CMC"
 FT 15..418
 FT /note- "ketosynthase domain, KS'(s)"
 FT 525..882
 FT /note- "acyltransferase domain, AT(s)"
 FT 942..1025
 FT /note- "acyl carrier protein domain, ACP(s)"
 FT 1060..1483
 FT /note- "ketosynthase domain, KSI"
 FT 1596..1953
 FT /note- "acyltransferase domain, AT1"
 FT 2232..2416
 FT /note- "ketoreductase domain, KR1"
 FT 2533..2616
 FT /note- "acyl carrier protein domain, ACP1"
 FT 2641..3064
 FT /note- "ketosynthase domain, KS2"
 FT 3184..3520
 FT /note- "acyltransferase domain, AT2"
 FT 3546..3727
 FT /note- "dehydratase domain, DH2"
 FT 4083..4268
 FT /note- "ketoreductase domain, KR2"
 FT 4374..4457
 FT /note- "acyl carrier protein domain, ACP2"
 FT Domain

XX EP791655-A2.
 PN
 XX 27-AUG-1997.
 PD
 XX 19-FEB-1997; 57EP-0301056.
 PF
 XX 22-FEB-1996; 96US-0012078.
 PR
 XX (ELIL) LILLY & CO ELI.
 PA
 PI Dehoff BS, Künstross SA, Rostreck PR, Sutton KL;
 DR WPI: 1997-418046/39.
 DR N-PSDB: T80414.
 DR
 PT DNA encoding Streptomyces fradiae ty lactone synthase domain - for
 PT production of tylosin-related polyketide compounds
 PT
 XX
 PS Example 2; Pages 135-149; 220pp; English.
 XX
 CC W22606-W22610 represent proteins encoded by the platenolide synthase gene
 CC cluster. The gene cluster is also referred to as the srmg gene, and was
 CC isolated from Streptomyces ambofaciens. These sequences are
 CC multi-functional proteins which direct the synthesis of the polyketide
 CC platenolide. Platenolide is the basic building block of the macroide
 CC antibiotic spiramycin. The DNA encoding this sequence was used along with
 CC the tylg gene (see T80413) to create a hybrid ORF1 sequence (see T80415).
 CC The tylg gene is the ty lactone synthase gene cluster of the invention.
 CC The tylg sequence was isolated from Streptomyces fradiae, and encodes
 CC multifunctional proteins which direct the synthesis of the polyketide
 CC ty lactone. Ty lactone is the basic building block of the antibiotic
 CC tylosin. The hybrid sequence can be used to transform S. ambofaciens
 CC lacking the srmg ORF1 sequence, or S. fradiae lacking the tylg ORF1
 CC sequence, so that they can produce polyketides. The DNA sequence can be
 CC modified so as to alter the type of carboxylic acids incorporated, the
 CC number of carboxylic acids incorporated and/or the post-condensation
 CC reactions performed, thereby resulting in novel tylosin-related
 CC polyketides.
 CC XX
 SQ Sequence 4550 AA;

Query Match 83.3%; Score 35; DB 18; Length 4550;
 Best Local Similarity 85.7%; Pred. No. 1.5e+03;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLGRPG 8
 Db 4223 gllgrpg 4229
 ||:|||||

RESULT 8
 G11597 G11597 standard; Protein; 499 AA.
 XX
 AC G11597;
 XX
 DT 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 10372.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 XX
 OS Arabidopsis thaliana.
 OS
 PN EP1033405-A2.
 PN
 PD 06-SEP-2000.
 PD
 XX 25-FEB-2000; 2000EP-0301439.
 XX

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 25-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 81.0%; Score 34; DB 21; Length 499;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGVLAGRP 7

Db 490 vgywgrp 496

RESULT 9

ID G52088 standard; Protein: 499 AA.

AC G52088;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 66175.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

OS Arabidopsis thaliana.

PN EPI03405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-01444325.

	PR	19-JUL-1999;	99US-0144331
	PR	19-JUL-1999;	99US-0144332
	PR	19-JUL-1999;	99US-0144333
	PR	19-JUL-1999;	99US-0144334
	PR	19-JUL-1999;	99US-0144335
	PR	20-JUL-1999;	99US-0144632
	PR	20-JUL-1999;	99US-0144887
	PR	21-JUL-1999;	99US-0144881
	PR	21-JUL-1999;	99US-0145086
	PR	22-JUL-1999;	99US-0145088
	PR	22-JUL-1999;	99US-0145087
	PR	22-JUL-1999;	99US-0145089
	PR	23-JUL-1999;	99US-0145112
	PR	23-JUL-1999;	99US-0145118
	PR	26-JUL-1999;	99US-0145218
	PR	27-JUL-1999;	99US-0145224
	PR	27-JUL-1999;	99US-0145276
	PR	28-JUL-1999;	99US-0145913
	PR	02-AUG-1999;	99US-0145918
	PR	02-AUG-1999;	99US-0145951
	PR	02-AUG-1999;	99US-0146386
	PR	03-AUG-1999;	99US-0146388
	PR	04-AUG-1999;	99US-0146389
	PR	04-AUG-1999;	99US-0147038
	PR	05-AUG-1999;	99US-0147204
	PR	05-AUG-1999;	99US-0147302
	PR	06-AUG-1999;	99US-0147192
	PR	06-AUG-1999;	99US-0147260
	PR	06-AUG-1999;	99US-0147303
	PR	09-AUG-1999;	99US-0147416
	PR	09-AUG-1999;	99US-0147493
	PR	11-AUG-1999;	99US-0147935
	PR	13-AUG-1999;	99US-0148177
	PR	13-AUG-1999;	99US-0148319
	PR	13-AUG-1999;	99US-0148341
	PR	16-AUG-1999;	99US-0148565
	PR	16-AUG-1999;	99US-0148684
	PR	17-AUG-1999;	99US-0149368
	PR	18-AUG-1999;	99US-0149375
	PR	20-AUG-1999;	99US-0149426
	PR	20-AUG-1999;	99US-0149742
	PR	20-AUG-1999;	99US-0149723
	PR	22-AUG-1999;	99US-0149929
	PR	22-AUG-1999;	99US-0149930
	PR	22-AUG-1999;	99US-0150566
	PR	27-AUG-1999;	99US-0150884
	PR	27-AUG-1999;	99US-0151065
	PR	30-AUG-1999;	99US-0151066
	PR	31-AUG-1999;	99US-0151080
	PR	01-SEP-1999;	99US-0151303
	PR	07-SEP-1999;	99US-0151308
	PR	10-SEP-1999;	99US-0151930
	PR	13-SEP-1999;	99US-0152363
	PR	15-SEP-1999;	99US-0153758
	PR	16-SEP-1999;	99US-0154018
	PR	20-SEP-1999;	99US-0154079
	PR	22-SEP-1999;	99US-0154779
	PR	23-SEP-1999;	99US-0155139
	PR	24-SEP-1999;	99US-0155486
	PR	28-SEP-1999;	99US-0156569
	PR	29-SEP-1999;	99US-0156658
	PR	04-OCT-1999;	99US-0157117
	PR	05-OCT-1999;	99US-0157753
	PR	07-OCT-1999;	99US-0157865
	PR	08-OCT-1999;	99US-0158029
	PR	12-OCT-1999;	99US-0158232
			99US-0158369

PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	14-OCT-1999	9905-0159684
PR	18-OCT-1999	9905-0160767
PR	18-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160769
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160814
PR	21-OCT-1999	9905-0160815
PR	21-OCT-1999	9905-0160980
PR	21-OCT-1999	9905-0160981
PR	21-OCT-1999	9905-0160982
PR	22-OCT-1999	9905-0161404
PR	22-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	25-OCT-1999	9905-0161407
PR	25-OCT-1999	9905-0161359
PR	26-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161361
PR	26-OCT-1999	9905-0161920
PR	28-OCT-1999	9905-0161992
PR	28-OCT-1999	9905-0161993
PR	29-OCT-1999	9905-0162142

Query Match	81.0%	Score 34	DB 21	Length 499
Best Local Similarity	85.7%	Pred. No. 2.7e+02		
Matches 6; Conservative	1	Mismatches 0	Indels 0	Gaps 0

27	1	vgvlgrr	/
Db	490	vgvmgrr	496

RESULT	10
G11596	
ID	G11596 standard; Protein; 564 AA
XX	

XX 17-OCT-2000 (first entry)
DT
XX
DE

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter termination sequence.

OS Arabidopsis thaliana.
XX

PN	EP1033405-A
XX	
PD	06-SEP-2000
XX	

AA 25-FEB-2000; 2000EP-0301439.
PF
XX

PR	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0123548.

PR	25-MAR-1999;	99US-0125788.
PR	29-MAR-1999;	99US-0126264.
PR	01-APR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.

PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 81.0%; Score 34; DB 21; Length 564;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VGVLRNP 7

DB 555 vgvmrp 561

RESULT 11

G52087 standard; Protein; 564 AA.

XX G52087;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 66174.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144336.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 81.0%; Score 34; DB 21; Length 564;
Best Local Similarity 85.7%; Pred. NO. 3,1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VGVIGRP 7
Db 555 ygvmgp 561

RESULT 12
G52086
ID G52086 standard; Protein: 616 AA.
XX AC G52086;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 66173.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS Arabidopsis thaliana.
XX PN Ep1033405-A2.
PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139465.
PR 18-JUN-1999; 99US-0139470.
PR 21-JUN-1999; 99US-0139753.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0141854.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 23-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 81.0%; Score 34; DB 21; Length 616;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VGVIGRP 7
Db 607 vgvimgp 613

RESULT 13
G11595
ID G11595 standard; Protein: 617 AA.
XX G11595;
XX G11595;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10370.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148655.
 PR 16-AUG-1999; 99US-0148684.
 PR 17-AUG-1999; 99US-0149368.
 PR 18-AUG-1999; 99US-0149175.
 PR 20-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 25-AUG-1999; 99US-0149930.
 PR 26-AUG-1999; 99US-0150566.
 PR 27-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151438.
 PR 10-SEP-1999; 99US-0151930.
 PR 13-SEP-1999; 99US-0152363.
 PR 15-SEP-1999; 99US-0153070.
 PR 15-SEP-1999; 99US-0153798.
 PR 16-SEP-1999; 99US-0154018.
 PR 20-SEP-1999; 99US-0154039.
 PR 22-SEP-1999; 99US-0154779.
 PR 23-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0156458.
 PR 04-OCT-1999; 99US-0156596.
 PR 05-OCT-1999; 99US-0157117.
 PR 06-OCT-1999; 99US-0157753.
 PR 07-OCT-1999; 99US-0157865.
 PR 08-OCT-1999; 99US-0158029.
 PR 12-OCT-1999; 99US-0158232.
 PR 13-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 21-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160711.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 81.0%; Score 34; DB 21; Length 617;
 Best Local Similarity 85.7%; Pred. No. 3.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GVLGRP 7
 DB 608 vgvmgrrp 614

RESULT 14
 ID W16303 standard; Protein; 785 AA.
 AC W16303;
 XX 07-AUG-1997 (first entry)

DE Yeast mitochondria COX1 gene intron a12 protein.
 XX Nucleotide integrase; DNA cleavage; COX1 gene intron a12.

OS Saccharomyces cerevisiae.
 XX W09710362-A1.

PD 20-MAR-1997.
 XX 11-SEP-1996; 96MO-US14609.

PR 12-SEP-1995; 95US-0526964.
 XX (OHIO-) OHIO STATE RES FOUND.

PI Guo H, Lambowitz AM, Yang J, Zimmerman S;
 DR WPI: 1997-202256/18.

DR N-PSDB; T63319.
 XX New ribonucleoprotein particle preparations - which can cleave

PT double- or single-stranded DNA and attach a nucleic acid molecule to
 the cleaved strand
 XX Disclosure: Page 40-42; 58pp; English.

PS Yeast mitochondrial COX1 gene intron a12-encoded protein (W16303)
 XX can be used in a novel method for cleaving a double-stranded or

CC single-stranded DNA substrate. The method utilizes a nucleotide
 CC integrase, pref. a (reconstituted) ribonucleoprotein (RNP) particle
 CC comprising an excised RNA transcript encoded by the group II intron

CC a12 (see also T63319) of the COX1 gene and a protein encoded by the
 CC group II intron a12 of the COX1 gene. The RNP is incubated with the
 CC DNA substrate in the presence of a divalent cation so that at least
 CC one strand of the DNA substrate is cleaved and a nucleic acid
 CC molecule is attached to one or both of the cleaved strands. The
 CC COX1 gene intron a11 RNA (see also T63320) and protein (W16304) can
 CC also be used.
 CC
 SQ Sequence 785 AA;

Query Match 81.0%; Score 34; DB 18; Length 785;
 Best Local Similarity 71.4%; Pred. No. 4.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GVLGRP 8

DB 534 g1l9kpg 540

RESULT 15
 Y35025
 ID Y35025 standard; Protein; 467 AA.

```

XX AC Y35025;
XX DT 13-SEP-1999 (first entry)
XX DE Chlamydia pneumoniae transmembrane protein sequence.
XX DE
XX DE Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX DE sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX DE vaccine; neutralising epitope.
XX OS Chlamydia pneumoniae.
XX PM W09927105-A2.
XX PD 03-JUN-1999.
XX PF 20-NOV-1998; 98WO-1B01890.
XX PR 04-NOV-1998; 98US-0107078.
XX PR 21-NOV-1997; 97ER-0014673.
XX PA (GEST ) GENSET.
XX PI
XX PI Griffiths R;
XX DR WPI: 1999-357842/30.
XX PT
XX PT Genome sequence of Chlamydia pneumoniae
XX PS
XX PS Page 921-922; Disclosure; 1912pp; English.
XX CC Y34584-Y35879 represent the proteins encoded by all the open reading
XX CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
XX CC C. pneumoniae causes respiratory disease such as pneumonia and
XX CC bronchitis and is thought to be a contributing factor in heart
XX CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
XX CC nodosum or pharyngitis. The polypeptides encoded by the open reading
XX CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
XX CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
XX CC nucleotide sequences can also be used as immunogenic compositions,
XX CC especially where the vector directs the expression of a neutralising
XX CC epitope of C. pneumoniae.
XX SO
XX SO Sequence 467 AA:

Query Match 78.6%; Score 33; DB 20; Length 467;
Best Local Similarity 62.5%; Pred. No. 3.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 VGVLCRPG 8
   ||: ||: ||
Db 156 vgiagkpg 163

```

Search completed: June 13, 2001, 14:25:48
Job time: 667 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:12 ; Search time 87.97 Seconds
(without alignments)
6.250 Million cell updates/sec

Title: PCT-US01-05825A-29

Perfect score: 42

Sequence: 1 VGVLGRG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues.

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_67:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	90.5	812	2	S31521 collagen COLF1 - f
2	36	85.7	485	2	G83359 hypothetical prote
3	36	85.7	577	2	B75585 probable long-chain
4	36	85.7	1616	2	I37183 gene APXL protein
5	35	83.3	317	2	T35010 probable integral
6	35	83.3	321	2	C70596 hypothetical prote
7	35	83.3	356	2	S15907 collagen alpha 1(I
8	35	83.3	411	2	E70667 hypothetical prote
9	35	83.3	1690	1	CGHU1B collagen alpha 4(I
10	35	83.3	3198	2	A43426 collagen alpha 2 f
11	34	81.0	283	2	T32921 hypothetical prote
12	34	81.0	294	2	T21668 hypothetical prote
13	34	81.0	324	2	T18763 hypothetical prote
14	34	81.0	324	2	A31920 collagen sqt-1 pre
15	34	81.0	382	2	S20375 collagen alpha 3(V
16	34	81.0	3643	2	T36410 probable polyketid
17	33	78.6	284	2	T29528 hypothetical prote
18	33	78.6	300	2	T24482 hypothetical prote
19	33	78.6	376	2	D64897 probable quinoe
20	33	78.6	450	2	B72081 replication initia
21	33	78.6	467	2	A81590 chromosomal replic
22	33	78.6	471	2	S15035 acetylcholinestera
23	33	78.6	600	2	S76764 hypothetical prote
24	33	78.6	744	2	S65669 biotin sulfoxide r
25	33	78.6	879	2	B70014 antibiotic synthe
26	33	78.6	1007	2	T47430 disease resistance
27	33	78.6	1633	2	JC5056 polybromo 1 - chic
28	33	78.6	1669	1	CGMS4B collagen alpha 1(I
29	33	78.6	1838	1	CGHU1V collagen alpha 1(V

30	33	78.6	1843	2	S18803 collagen alpha 1(V
31	33	78.6	2129	2	T14182 fxdc protein - Myc
32	33	78.6	2297	2	T34918 polyketide synthas
33	33	78.6	2512	1	MMWVS nonstructural poly
34	33	78.6	2514	1	MMWVS nonstructural poly
35	33	78.6	2569	2	T14164 peptidase synthetase
36	33	78.6	2944	2	A54849 collagen alpha 1(V
37	33	78.6	4077	2	T17484 hypothetical prote
38	33	78.6	4735	2	T17463 rifamycin polyket
39	33	78.6	5149	2	F83345 probable non-ribos
40	33	78.6	26926	1	I38344 titin, cardiac mus
41	32	76.2	161	2	F72593 hypothetical prote
42	32	76.2	178	1	R5RT11 ribosomal protein
43	32	76.2	178	2	S45049 ribosomal protein
44	32	76.2	184	2	S60245 ribosomal protein
45	32	76.2	203	2	A72705 hypothetical prote

ALIGNMENTS

RESULT 1

S31521

collagen COLF1 - freshwater sponge (Ephydatia muelleri)

C:Species: Ephydatia muelleri

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999

C:Accession: S31521

R:Exposito, J.Y.: van der Rest, M.; Garrone, R. submitted to the EMBL Data Library, December 1992

A:Description: The complete intron/exon structure of E. muelleri collagen gene sugges

A:Reference number: S31521

A:Accession: S31521

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-812 <EXP>

A:Cross-references: EMBL:X69818; NID:9429036; PIDN:CAA49472.1; PID:99300

C:Superfamily: unassigned collagens

Query Match 90.5%; Score 38; DB 2; Length 812;
Best Local Similarity 87.5%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VGVLGRG 8

DB 203 VGADGRPG 210

RESULT 2

G83359

hypothetical protein PA2283 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: G83359

S:Stover, C.K.; Pham, X.O.; Ewlin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; adman, S.; Van, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L. Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; M0ID:20437337

A:Accession: G83359

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-485 <STO>

A:Cross-references: GB:AE004654; GB:AE004091; NID:99948311; PIDN:AAG05671.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2283

Query Match 85.7%; Score 36; DB 2; Length 485;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGVLRPG 8
 ||:||||
 Db 336 VGLGKPG 343

RESULT 3

probable long-chain-fatty-acid-CoA ligase (EC 6.2.1.3) DRA0309 [similarity] - Deinococcus
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
 C:Accession: B75585
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: B75585
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-577 <WHI>
 A:Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDN:AAF12469.1; PID:9646076
 C:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0309
 A:Map position: 2
 C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
 C:Keywords: acid-thiol ligase; coenzyme A
 F:85-561/Domain: acetate-CoA ligase homology <ACLD>

Query Match
 Best Local Similarity 85.7%; Score 36; DB 2; Length 577;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGVLRPG 7
 ||:||||
 Db 495 VGLGKPG 501

RESULT 4

gene APXL protein - human
 C:Species: Homo sapiens (man)
 C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
 C:Accession: I37183
 R:Schlaifflin, M.V.; Bassil, M.T.; Rugarli, E.L.; Renteria, A.; Galli, L.; Ballabio, A.
 Hum. Mol. Genet. 4, 373-382, 1995
 A:Title: Cloning of a human homologue of the Xenopus laevis APX gene from the ocular alh
 A:Reference number: I37183; MUID:95315933
 A:Accession: I37183
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1616 <RES>
 A:Cross-references: EMBL:X83543; NID:9790999; PIDN:CAA58534.1; PID:g1181628
 C:Genetics:
 A:Gene: GDB:APXL
 A:Cross-references: GDB:582527
 A:Map position: Xp22.3-xp22.3
 C:Superfamily: G1GF domain homology
 F:32-104/Domain: G1GF domain homology <G1G>

Query Match
 Best Local Similarity 85.7%; Score 36; DB 2; Length 1616;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VGVLRPG 7
 ||:||||
 Db 1087 VGVLRPG 1093

RESULT 5

probable integral membrane protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C:Accession: T35010
 R:Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A:Reference number: 221565
 A:Accession: T35010
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-317 <SE>
 A:Cross-references: EMBL:AL079355; PIDN:CAB45566.1; GSPDB:GN00070; SCODEB:SC4C6.02C
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCODEB:SC4C6.02C

Query Match
 Best Local Similarity 83.3%; Score 35; DB 2; Length 317;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVLRPG 8
 ||:||||
 Db 142 IGVLSPG 149

RESULT 6

hypothetical protein RV3218 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: C70596
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 Rajandream, M.A.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Solstun, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A:Reference number: A70500; MUID:98295987
 A:Accession: C70596
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-321 <COI>
 A:Cross-references: GB:Z95120; GB:AL123456; NID:93261739; PIDN:CAB08319.1; PID:g20726
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3218
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV3218

Query Match
 Best Local Similarity 83.3%; Score 35; DB 2; Length 321;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VGLRPG 8
 ||:||||
 Db 73 GMLGKPG 79

RESULT 7

collagen alpha 1(IV) chain - bovine (fragments)
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 04-Dec-1997 #sequence_revision 05-Apr-1995 #text_change 19-Oct-1995
 C:Accession: A39474; S16907; S18432; A39419
 R:Gunwar, S.; Noelken, M.E.; Hudson, B.G.
 J. Biol. Chem. 266, 14088-14094, 1991
 A:Title: Properties of the collagenous domain of the alpha3(IV) chain, the Goodpastur
 le helical structure and noncollagenous domain.
 A:Reference number: A39474; MUID:91310700
 A:Accession: A39474

A:Molecule type: protein
 A:Residues: 1-9 <GU2>
 R:Schuppan, D.; Glaville, R.W.; Timpl, R.; Dixit, S.N.; Kang, A.H.
 Biochem. J. 220, 227-233, 1984
 A:Title: Sequence comparison of pepsin-resistant segments of basement-membrane collagen
 A:Reference number: S16907; PMID:84256630
 A:Accession: S16907
 A:Molecule type: protein
 A:Residues: 4-8, E, 10-68; 69-158; 159-321; 322-349 <SC2>
 R:Butkowski, R.J.; Langeveld, J.P.M.; Wierslander, J.; Hamilton, J.; Hudson, B.G.
 J. Biol. Chem. 262, 7874-7877, 1987
 A:Title: Localization of the Goodpasture epitope to a novel chain of basement membrane C
 A:Reference number: S18432; PMID:87222419
 A:Accession: S18432
 A:Molecule type: protein
 A:Residues: 337-347, 'FL', 350-353 <BU2>
 R:Gunnar, S.; Ballesater, F.; Kalluri, R.; Timoneda, J.; Chonko, A.M.; Edwards, S.J.; Nog
 J. Biol. Chem. 266, 15318-15324, 1991
 A:Title: Glomerular basement membrane. Identification of dimeric subunits of the noncoll
 A:Reference number: A39419; PMID:91332055
 A:Accession: A39419
 A:Molecule type: protein
 A:Residues: 337-347, 'FL', 350-356 <GU3>
 C:Superfamily: collagen alpha 1(IV) chain
 C:Keywords: basement membrane; cell binding; coiled coll.; disulfide bond; extracellular
 F:31,34,37,46,61,69,78,84,87,102,110,122,125,137,140,143,149,155,158,161,164,185,188,195
 proline (Pro) #status experimental

Query Match 83.3%; Score 35; DB 2; Length 356;
 Best Local Similarity 62.5%; Pred. No. 41;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVLRPG 8
 : : : : :
 DB 104 IGXGRPG 111

RESULT 8
 E70667
 hypothetical protein RY1869c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: E70667
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; PMID:98295987
 A:Accession: E70667
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-411 <COL>
 A:Cross-references: GB:283859; GB:AL123456; NID:q3261678; PIDN:CAB06118.1; PID:91781178
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RY1869c
 C:superfamily: collagen dioxigenase ferredoxin reductase component

Query Match 83.3%; Score 35; DB 2; Length 411;
 Best Local Similarity 85.7%; Pred. No. 47;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLGRPG 8
 : : : : :
 DB 311 GMLGRPG 317

CGHUB

RESULT 9

CGHUB

collagen alpha 4(IV) chain precursor - human
 N:Alternate names: procollagen alpha 4(IV) chain
 C:Species: Homo sapiens (man)
 C:Date: 06-Feb-1995 #sequence_revision 03-Oct-1995 #text_change 16-Jun-2000
 R:Leinonen, A.; Maryama, M.; Mochizuki, T.; Tytgvason, K.; Reiders, S.T.
 J. Biol. Chem. 269, 26172-26177, 1994
 A:Title: Complete primary structure of the human type IV collagen alpha4(IV) chain. C
 A:Reference number: A35360; PMID:95014445
 A:Accession: A35360
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1690 <LE1>
 A:Cross-references: GB:X81053; NID:q574805; PIDN:CA56943.1; PID:q574806
 R:Sugimoto, M.; Ohashi, T.; Yoshioke, H.; Matsuo, N.; Ninomiya, Y.
 FEBS Lett. 330, 122-128, 1993
 A:Title: cDNA isolation and partial gene structure of the human alpha-4(IV) collagen
 A:Reference number: S36854; PMID:93374047
 A:Accession: S36854
 A:Molecule type: DNA; mRNA
 A:Residues: 1219-1658, 'FE', 1661-1690 <SU0>
 A:Cross-references: DDBJ:D17391; NID:q440365; PIDN:BA04214.1; PID:q457161
 A:Experimental source: whole eye
 R:Kamagata, Y.; Matel, M.G.; Ninomiya, Y.
 J. Biol. Chem. 267, 23753-23758, 1992
 A:Title: Isolation and sequencing of cDNAs and genomic DNAs encoding the alpha4 chain
 A:Reference number: S28777; PMID:93054733
 A:Accession: S28777
 A:Molecule type: DNA
 A:Residues: 1407-1424, 'G', 1426-1430, 'A', 1432-1439, 'L', 1441-1507 <KAM>
 A:Cross-references: GB:101475; GB:101476
 A:Note: the codons given for 1438-Asp (GAG) and 1443-gly (GCA) are inconsistent with
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 ed and subsequently O-glycosylated.
 C:Genetics:
 A:Gene: GDB:COL4A4
 A:Cross-references: GDB:132673; OMIM:120131
 A:Map position: 2q35-2q37
 A:Introns: 39/1; 1406/1; 1445/1; 1508/1; 1603/3 #status incomplete
 A:Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands w
 C:Complex: this minor type IV collagen is thought to form a heterotrimer of two alpha
 mong trimer amino-terminal domains (with disulfide and desmosine cross-links), dimer
 er associations in the interrupted helical domain (with disulfide and desmosine cross
 C:Function:
 A:Description: minor structural component of extracellular basement membrane in kidne
 C:Superfamily: collagen alpha 1(IV) chain
 C:Keywords: basement membrane; coiled coll.; extracellular matrix; glycoprotein; hydro
 F:1-38/Domain: signal sequence #status predicted <SIG>
 F:39-1690/Product: collagen alpha 4(IV) chain #status predicted <MAT>
 F:39-61/Domain: amino-terminal nonhelical, NH1 <NH1>
 F:62-1466/Region: interrupted helical
 F:94-96/Region: cell attachment (R-G-D) motif
 F:145-147/Region: cell attachment (R-G-D) motif
 F:189-191/Region: cell attachment (R-G-D) motif
 F:310-312/Region: cell attachment (R-G-D) motif
 F:724-726/Region: cell attachment (R-G-D) motif
 F:785-787/Region: cell attachment (R-G-D) motif
 F:989-991/Region: cell attachment (R-G-D) motif
 F:1212-1214/Region: cell attachment (R-G-D) motif
 F:1467-1690/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
 F:1471-1569/Domain: collagen IV carboxyl-terminal repeat <CT1>
 F:1579-1686/Domain: collagen IV carboxyl-terminal repeat <CT2>
 F:47,52,55,57,266,400,460,492,494,668,790,828,1095,1131,1294,1317,1375,1407/Disulfide
 F:142,669/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:1480-1566,1513-1569/Disulfide bonds: (or 1480-1569, 1513-1566) #status predicted
 F:1525-1531,1634-1641/Disulfide bonds: #status predicted
 F:1588-1683,1622-1686/Disulfide bonds: (or 1588-1686, 1622-1683) #status predicted

Query Match 83.3%; Score 35; DB 1; Length 1690;
 Best Local Similarity 85.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLAGRPG 8
 Db 391 GLLAGRPG 397

RESULT 10
 A43426

collagen alpha 2 fibrillar chain precursor - sea urchin (Strongylocentrotus purpuratus)
 C:Species: Strongylocentrotus purpuratus (purple urchin)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: A43426
 R:Exposito, J.Y.; D'Alessio, M.; Ramirez, F.
 J. Biol. Chem. 267, 17404-17408, 1992
 A:Title: Novel amino-terminal propeptide configuration in a fibrillar procollagen underg
 A:Reference number: A43426; MUID:92381062
 A:Accession: A43426
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-3198 <EXP>
 A:Cross-references: GB:M92041; NID:g161448; PIDN:AAA30040.1; PID:g161449
 A:Note: sequence extracted from NCBI backbone (NCBIP:111965)
 C:Superfamily: unassigned collagens: fibrillar collagen carboxyl-terminal homology; von
 F;2978-3198/Domain: fibrillar collagen carboxyl-terminal homology <RCC>

Query Match 83.3%; Score 35; DB 2; Length 3198;
 Best Local Similarity 75.0%; Pred. No. 3.3e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VGVLAGRPG 8
 Db 1986 VGISGRPG 1993

RESULT 11
 T32921

hypothetical protein K09H9.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T32921
 R:Madsen, C.; Graves, T.; Blair, T.
 submitted to the EMBL Data Library, January 1998
 A:Description: The sequence of C. elegans cosmid K09H9.
 A:Reference number: Z21247
 A:Accession: T32921
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-283 <MAD>
 A:Cross-references: EMBL:AF043700; PIDN:AAB9570.1; GSPDB:GN00019; CESP:K09H9.3
 A:Experimental source: strain Bristol N2; clone K09H9
 C:Genetics:
 A:Gene: CESP:K09H9.3
 A:Map position: 1
 A:Introns: 87/1
 C:Superfamily: unassigned collagens

Query Match 81.0%; Score 34; DB 2; Length 283;
 Best Local Similarity 75.0%; Pred. No. 50;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VGVLAGRPG 8
 Db 209 VGLAGRPG 216

RESULT 12
 T21668

hypothetical protein F32G8.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T21668

R:McMurray, A.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19456
 A:Accession: T21668
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-294 <WIL>
 A:Cross-references: EMBL:272509; PIDN:CAA96649.1; GSPDB:GN00023; CESP:F32G8.5
 A:Experimental source: clone F32G8
 C:Genetics:
 A:Gene: CESP:F32G8.5
 A:Map position: 5
 A:Introns: 47/3
 C:Superfamily: unassigned collagens

Query Match 81.0%; Score 34; DB 2; Length 294;
 Best Local Similarity 75.0%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VGVLAGRPG 8
 Db 156 VGINGRPG 163

RESULT 13
 T18763

hypothetical protein B0491.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T18763
 R:Sulston, J.
 submitted to the EMBL Data Library, June 1995
 A:Reference number: Z19019
 A:Accession: T18763
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-324 <WIL>
 A:Cross-references: EMBL:249907; PIDN:CMA90084.1; GSPDB:GN00020; CESP:B0491.2
 A:Experimental source: clone B0491
 C:Genetics:
 A:Gene: CESP:B0491.2
 A:Map position: 2
 A:Introns: 47/3
 C:Superfamily: unassigned collagens

Query Match 81.0%; Score 34; DB 2; Length 324;
 Best Local Similarity 85.7%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 2 GVLAGRPG 8
 Db 178 GALGRPG 184

RESULT 14
 A31920

collagen sgt-1 precursor - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 20-Sep-1999
 C:Accession: A31920
 R:Kramer, J.M.; Johnson, J.J.; Edgar, R.S.; Basch, C.; Roberts, S.
 Cell 55, 555-565, 1988
 A:Title: The sgt-1 gene of C. elegans encodes a collagen critical for organismal morp
 A:Reference number: A31920; MUID:89028667
 A:Accession: A31920
 A:Molecule type: DNA
 A:Residues: 1-324 <KRA>
 A:Cross-references: GB:J03146; NID:g156445; PIDN:AAA65468.1; PID:g156446
 C:Genetics:
 A:Introns: 47/3
 C:Superfamily: unassigned collagens

Query Match 81.0%; Score 34; DB 2; Length 324;
Best Local Similarity 85.7%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVLCRPG 8
| | | | |
DB 178 GVLGRPG 184

RESULT 15

S20375
collagen alpha 3(V) chain - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-May-1997
C:Accession: S20375
R:Mann, K.
Biol. Chem. Hoppe-Seyler 373, 69-75, 1992
A:Title: Isolation of the alpha-3-chain of human type V collagen and characterization by
A:Reference number: S20375; MUID:92239022
A:Accession: S20375
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-382 <MAN>
C:Superfamily: unassigned collagens

Query Match 81.0%; Score 34; DB 2; Length 382;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVLCRPG 8
| | | | |
DB 6 VGLTGRPG 13

Search completed: June 13, 2001, 14:23:12
Job time: 743 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:30:37 ; Search time 51.57 Seconds
(without alignments)
5.314 Million cell updates/sec

Title: PCT-US01-05825A-29

Perfect score: 42

Sequence: 1 VGVLAGRPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	85.7	1616	1 APXL_HUMAN	Q13796 homo sapien
2	35	83.3	1690	1 CA44_HUMAN	P53420 homo sapien
3	34	81.0	324	1 CCS1_CAEEL	P12114 caenorhabdi
4	34	81.0	382	1 CA35_HUMAN	P25940 homo sapien
5	33	78.6	284	1 YKNP_CAEEL	P21285 caenorhabdi
6	33	78.6	353	1 YNCB_ECOLI	P76113 escherichia
7	33	78.6	471	1 COLO_TORMA	O03637 torpedo mar
8	33	78.6	744	1 BISC_RHOSE	P54934 rhodobacter
9	33	78.6	1669	1 CA14_MOUSE	P02463 mus musculu
10	33	78.6	1838	1 CA15_HUMAN	P20908 homo sapien
11	33	78.6	2512	1 POLN_SINDV	P03317 sindbis vir
12	33	78.6	2514	1 POLN_SINDO	P27283 sindbis vir
13	33	78.6	2944	1 CA17_HUMAN	Q02388 homo sapien
14	32	76.2	165	1 RUL1_PIG	O29205 sus scrofa
15	32	76.2	177	1 RUL1_HUMAN	P39026 homo sapien
16	32	76.2	184	1 RUL1_DROME	P46222 drosophila
17	32	76.2	215	1 RIBB_BUCAT	P57167 buchnera ap
18	32	76.2	245	1 NODG_RHIS3	P27337 rhizobium m
19	32	76.2	397	1 PGK_ZYMO	P09404 zygomonas m
20	32	76.2	456	1 COLO_HUMAN	O39215 homo sapien
21	32	76.2	501	1 YH92_CAEEL	O23256 caenorhabdi
22	32	76.2	681	1 NU5C_STYNY3	O55432 synechocyst
23	32	76.2	720	1 KREG_YEAST	P22486 saccharomyc
24	32	76.2	737	1 SKN1_CANAL	P87023 candida alb
25	32	76.2	740	1 KREG_CANAL	P87023 candida alb
26	32	76.2	760	1 RELA_CORGL	O87331 corynebacte
27	32	76.2	771	1 SKN1_YEAST	P33336 saccharomyc
28	32	76.2	1333	1 YN99_YEAST	P33756 saccharomyc
29	32	76.2	1366	1 CA21_CANFA	O46392 canis famli
30	32	76.2	1490	1 CDR4_CANAL	O74676 candida alb
31	32	76.2	1499	1 CDR2_CANAL	P78595 candida alb
32	32	76.2	1501	1 CDR1_CANAL	P43071 candida alb
33	32	76.2	1501	1 CDR3_CANAL	O42690 candida alb

34	32	76.2	1501	1 SNO2_YEAST	P32568 saccharomyc
35	32	76.2	1511	1 PDR3_YEAST	P33302 saccharomyc
36	32	76.2	1529	1 PDRE_YEAST	O04182 saccharomyc
37	32	76.2	1564	1 PDRA_YEAST	P51533 saccharomyc
38	32	76.2	1775	1 CA14_DROME	P08120 drosophila
39	32	76.2	1806	1 CA1B_HUMAN	P12107 homo sapien
40	32	76.2	4687	1 PLE1_RAT	P30427 rattus norv
41	31	73.8	165	1 Y814_ARCFU	O29444 archaeglob
42	31	73.8	216	1 CTRD_NEIMA	P57013 neisseria m
43	31	73.8	216	1 CTRD_NEIMA	P32016 neisseria m
44	31	73.8	217	1 BEXA_HAEIN	P10640 haemophilus
45	31	73.8	236	1 COLO_MOUSE	O35348 mus musculu

ALIGNMENTS

```

RESULT 1
ID APXL_HUMAN STANDARD: PRT: 1616 AA.
AC Q13796;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE APICAL-LIKE PROTEIN (APXL PROTEIN).
GN APXL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=95315933; PubMed=7795590;
RA Schaffner V.M., Baesi M.T., Rugaril E.I., Renieri A., Galli L.,
RA Ballabio A.;
RT "Cloning of a human homologue of the Xenopus laevis APX gene from the
RT ocular albinism type 1 critical region."
RL Hum. Mol. Genet. 4:373-382(1995).
RN [2]
RP SEQUENCE OF 56-1616 FROM N.A.
RA Shen Y., Gibbs R.A.;
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
CC -1- TISSUE SPECIFICITY: ABUNDANT IN RETINA AND MELANOMA; ALSO IN
CC BRAIN, PLACENTA, LUNG, KIDNEY, PANCREAS.
CC -1- SIMILARITY: SOME, TO XENOPUS LAEVIS APICAL PROTEIN (APX).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X83543; CA58534.1; -
DR EMBL: AC002365; AAC32592.1; -
DR HSSP: Q12959; 1PDR.
DR MIM: 300103; -
DR InterPro: IPR000867; -
DR InterPro: IPR001478; -
DR Pfam: PF00219; IGFBP_1.
DR Pfam: PF00595; PDZ_1.
DR PROSITE: PS50106; PDZ_1.
FT SIMILAR 54 112
FT DOMAIN 150 153 TO MOUSE SYNTROPHIN-1 AND RABBIT
FT DOMAIN 314 320 DYSTROPHIN-ASSOCIATED PROTEIN.
FT DOMAIN 343 346 POLY-SER.
FT DOMAIN 1065 1068 POLY-PRO.
FT DOMAIN 1616 AA; 176409 MW; 7524065BC0B60A2 CRC64;
SO SEQUENCE

```

RESULT 2
CA44_HUMAN
ID CA44_HUMAN STANDARD; PRT; 1690 AA.
DT P53420;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE COLLAGEN ALPHA 4(IV) CHAIN PRECURSOR.
GN COL4A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RI [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95014445; PubMed=7523402;
RA Leinonen A., Marlyama M., Mochizuki T., Tryggvason K., Reenders S.T.;
RT "Complete primary structure of the human type IV collagen alpha 4(IV)
RT chain. Comparison with structure and expression of the other alpha
RT (IV) chains.";
RN J. Biol. Chem. 269:26172-26177(1994).
RN [2]
RP SEQUENCE OF 1219-1690 FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=93374047; PubMed=8365481;
RA Sugimoto M., Ohashi T., Yoshitaka H., Matsuo N., Ninomiya Y.;
RT "cDNA isolation and partial gene structure of the human alpha 4(IV)
RT collagen chain.";
RN FEBS Lett. 330:122-128(1993).
RN [3]
RP SEQUENCE OF 1407-1507 FROM N.A.
RX MEDLINE=93054733; PubMed=1429714;
RA Kamagata Y., Mattei M.-G., Ninomiya Y.;
RT "Isolation and sequencing of cDNAs and genomic DNAs encoding the
RT alpha 4 chain of basement membrane collagen type IV and assignment of
RT the gene to the distal long arm of human chromosome 2.";
RN J. Biol. Chem. 267:23753-23758(1992).
RN [4]
RP REVIEW ON VARIANTS.
RX MEDLINE=97338662; PubMed=9195222;
RA Lemmink H.H., Schroeder C.H., Monnens L.A.H., Smeets H.J.M.;
RT "The clinical spectrum of type IV collagen mutations.";
RN Hum. Mutat. 9:477-499(1997).
RN [5]
RP VARIANT AS SER-1201.
RX MEDLINE=95078927; PubMed=7987396;
RA Mochizuki T., Lemmink H.H., Matiyama M., Antignac C., Gubler M.-C.,
RA Plerson Y., Verellen-Dumoulin C., Chan B., Schroeder C.H.,
RA Smeets H.J., Reenders S.T.;
RT "Identification of mutations in the alpha 3(IV) and alpha 4(IV)
RT collagen genes in autosomal recessive Alport syndrome.";
RN Nat. Genet. 8:77-82(1994).
RN [6]
RP VARIANT BPH GIU-897.
RX MEDLINE=96379660; PubMed=8787673;
RA Lemmink H.H., Nillesen W.N., Mochizuki T., Schroeder C.H.,
RA Brunner H.G., Van Oost B.A., Monnens L.A.H., Smeets H.J.M.;
RT "Benign familial hematuria due to mutation of the type IV collagen
RT alpha4 gene.";
RN J. Clin. Invest. 98:1114-1118(1996).
RN [7]
RP VARIANTS AS, AND VARIANTS.
RX MEDLINE=99011253; PubMed=9792860;

RA Boye E., Mollet G., Forestier L., Cohen-Solal L., Heidet L.,
RT Cochot P., Gruenfeld J.-P., Palcoux J.-B., Gubler M.-C., Antignac C.;
RL "Determination of the genomic structure of the COL4A4 gene and of
RT novel mutations causing autosomal recessive Alport syndrome.";
CC Am. J. Hum. Genet. 63:1329-1340(1998).
CC -1- FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
CC GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
CC MESHWORX TOGETHER WITH LAMININS, PROTEGLYCANS AND ENCTACTIN/
CC NITROGEN.
CC
CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
CC
CC -1- TISSUE/CELLULAR LOCATION: CELL SURFACE (POTENTIAL).
CC
CC -1- SUBCELL SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE
CC COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE,
CC COCHLEA, LUNG AND BRAIN.
CC
CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
CC TRIPLE-HELICAL 7S DOMAIN.
CC
CC -1- PM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC
CC -1- PM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
CC IV COLLAGENS.
CC
CC -1- DISEASE: DEFECTS IN COL4A4 ARE ASSOCIATED WITH THE TYPE II
CC AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME (AS), AN HEREDITARY
CC GLOMERULOEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE,
CC HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN
CC MALES AND FEMALES.
CC
CC -1- DISEASE: DEFECTS IN COL4A4 ARE ASSOCIATED WITH FAMILIAL BENIGN
CC HEMATURIA (FBH) OR THIN BASEMENT MEMBRANE DISEASE. FBH IS
CC CHARACTERIZED BY PERSISTENT HEMATURIA, AN ELECTRON MICROSCOPICALLY
CC DETECTABLE THIN GLOMERULAR BASEMENT MEMBRANE (GBM) AND AN
CC AUTOSOMAL DOMINANT MODE OF INHERITANCE. RENAL FUNCTION REMAINS
CC NORMAL. IN CHILDREN DIFFERENTIATION BETWEEN FBH AND AS CAN BE
CC DIFFICULT, BECAUSE BOTH DISORDERS ARE MANIFESTED BY PERSISTENT
CC HEMATURIA AND THIN GBM AT THAT AGE.
CC
CC -1- SIMILARITY: TO OTHER TYPE IV COLLAGENS.

CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb-sib.ch/announce/>)
CC OR SEND AN EMAIL TO license@isb-sib.ch.
CC -----

DR EMBL: X81053; CAA56943.1; -;
DR EMBL: D17391; BAA04214.1; -;
DR MIM: 120131; -;
DR MIM: 141200; -;
DR MIM: 203780; -;
DR InterPro: IPR000087; -;
DR InterPro: IPR001442; -;
DR Pfam: PF01413; Cl: 2;
DR Pfam: PF01391; Collagen; 20.
KW Extracellular matrix; Connective tissue; Basement membrane; Repeat;
KW Hydroxylation; Collagen; Glycoprotein; Signal; Disease mutation;
KW Polymorphism.

FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1690 COLLAGEN ALPHA 4(IV) CHAIN.
FT DOMAIN 39 64 7S DOMAIN.
FT DOMAIN 65 1459 TRIPLE-HELICAL REGION.
FT DOMAIN 1460 1690 NONHELIICAL REGION (NC1).
FT SITE 94 96 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 145 147 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 189 191 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 310 312 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 724 726 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 785 787 CELL ATTACHMENT SITE (POTENTIAL).

```

FT SITE 989 991 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1206 1206 CLEAVAGE (BY COLLAGENASE)
FT SITE 1212 1214 (BY SIMILARITY).
FT DISULFID 1480 1566 OR 1566 (BY SIMILARITY).
FT DISULFID 1513 1566 OR 1566 (BY SIMILARITY).
FT DISULFID 1525 1531 BY SIMILARITY.
FT DISULFID 1588 1586 OR 1683 (BY SIMILARITY).
FT DISULFID 1622 1683 OR 1686 (BY SIMILARITY).
FT DISULFID 1634 1641 BY SIMILARITY.
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 441 446 MISSING (IN AS).
FT VARIANT 545 545 /FTID=VAR_008148.
FT VARIANT 570 570 G -> A.
FT VARIANT 897 897 /FTID=VAR_008149.
FT VARIANT 931 931 E -> O.
FT VARIANT 931 931 /FTID=VAR_008150.
FT VARIANT 931 931 G -> E (IN FBH).
FT VARIANT 1004 1004 /FTID=VAR_001912.
FT VARIANT 1004 1004 A -> T.
FT VARIANT 1030 1030 /FTID=VAR_008151.
FT VARIANT 1201 1201 L -> P.
FT VARIANT 1201 1201 /FTID=VAR_008152.
FT VARIANT 1402 1402 G -> V (IN AS).
FT VARIANT 1402 1402 /FTID=VAR_008153.
FT VARIANT 1572 1572 G -> S (IN AS).
FT VARIANT 1572 1572 /FTID=VAR_001913.
FT VARIANT 1572 1572 P -> S.
FT VARIANT 1572 1572 /FTID=VAR_008154.
FT VARIANT 1572 1572 P -> L (IN AS).
FT VARIANT 1572 1572 /FTID=VAR_008155.
FT CONFLICT 1659 1660 LQ -> FE (IN REF. 2).
FT SEQUENCE 1690 AA; 164095 MW; E1E72F283A72BAAE CRC64;

```

Query Match 83.3%; Score 35; DB 1; Length 1690;
 Best Local Similarity 85.7%; Pred. No. 1,1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLGRP 8
 DB 391 GVLGRP 397

```

RESULT 3
CCSI_CAEEL STANDARD: PRT; 324 AA.
ID CCSI_CAEEL 017509:
AC P1214; 017509:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CUTICLE COLLAGEN SQT-1.
GN SQT-1 OR ROL-5 OR B0491.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=89028667; PubMed=3180220;
RA Krimer J.M., Johnson J.J., Edgar R.S., Basch C., Roberts S.;
RT "The sqt-1 gene of C. elegans encodes a collagen critical for
RT organismal morphogenesis."
RL Cell 55:555-565(1988).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A

```

```

CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
CC -1- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
CC CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
CC LINKS.
CC -1- DISEASE: THIS IS A COLLAGEN CRITICAL FOR ORGANISMAL MORPHOGENESIS.
CC MUTATIONS IN SQT-1 CAN LENGTHEN, SHORTEN, OR HELICALLY TWIST THE
CC ENTIRE ANIMAL.
CC -1- SIMILARITY: TO OTHER COLLAGENS. STRONG, TO OTHER CUTICLE
CC COLLAGENS. ROL-6 AND SQT-1 BELONGS TO THE SAME GROUP OF COLLAGEN
CC AND MAY ALSO PHYSICALLY INTERACT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03146; AAA65468.1; -
DR EMBL: Z49907; CA930084.1; -
DR PIR: A31920; A31920.
DR WormPep: B0491.2; CE02104.
DR InterPro: IPR000087; -
DR Pfam: PF01391; Collagen: 2.
KW Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
FT DOMAIN 127 153
FT DOMAIN 171 231 TRIPLE-HELICAL REGION.
FT DOMAIN 237 299 TRIPLE-HELICAL REGION.
FT CONFLICT 158 158 V -> A (IN REF. 2).
FT CONFLICT 238 238 G -> R (IN REF. 2).
FT SEQUENCE 324 AA; 32779 MW; DBAC00826993C1CFC CRC64;

```

Query Match 81.0%; Score 34; DB 1; Length 324;
 Best Local Similarity 85.7%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVLGRP 8
 DB 178 GALGRP 184

```

RESULT 4
CA35_HUMAN STANDARD: PRT; 382 AA.
ID CA35_HUMAN P25940:
AC P25940:
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE COLLAGEN ALPHA 3(V) CHAIN (FRAGMENTS).
GN COL3A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1];
RP TISSUE=Placenta;
RC MEDLINE=92239022; PubMed=1571108;
RA Mann K.;
RT "Isolation of the alpha 3-chain of human type V collagen and
RT characterization by partial sequencing."
RL Biol. Chem. Hoppe-Seyler 373:69-75(1992).
CC -1- FUNCTION: TYPE V COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC (FIBRILLAR FORMING COLLAGEN). IT IS A MINOR CONNECTIVE TISSUE
CC COMPONENT OF NEARLY UBQUITOUS DISTRIBUTION. TYPE V COLLAGEN BINDS
CC TO DNA, HEPARAN SULFATE, THROMBOSPONDIN, HEPARIN, AND INSULIN.
CC -1- SUBUNIT: TRIMERS OF TWO ALPHA 1(V) AND ONE ALPHA 2(V) CHAINS IN
CC MOST TISSUES AND TRIMERS OF ONE ALPHA 1(V), ONE ALPHA 2(V), AND
CC ONE ALPHA 3(V) CHAINS IN PLACENTA.
CC -1- PTM: POLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.

```

CC -1- SIMILARITY: HIGH, TO ALPHA 1(V) AND ALPHA 1(XI) CHAINS.
 DR PIR: S20375; S20375.
 DR MIM: 120216; -
 DR Pfam: PF01391; Collagen; 6.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 FT NON-TER.
 FT NON-CONS 1 1
 FT NON-CONS 86 87
 FT NON-CONS 131 132
 FT NON-CONS 167 168
 FT NON-CONS 200 201
 FT NON-CONS 284 285
 FT NON-CONS 317 318
 FT NON-CONS 359 360
 FT NON-TER 382
 SO SEQUENCE 382 AA; 35352 MW; F8DBB5968F69B5C CRC64;

Query Match
 Best Local Similarity 81.0%; Score 34; DB 1; Length 382;
 Matches 6; Conservative 1; Pred. No. 40; Mismatches 1; Indels 0; Gaps 0;

QY 1 VGVLRPG 8
 Db 6 VGVLRPG 13

RESULT 5
 YKMP_CAEEL
 ID YKMP_CAEEL STANDARD; PRT; 284 AA.
 AC P91285;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE PUTATIVE CUTICLE COLLAGEN F27C1.8.
 GN F27C1.8.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Pelodermidae; Caenorhabditis.
 RN NCBI_TaxID=6239;
 RA
 RC STRAIN-BRISTOL N2;
 RL Wu X., Le T.T.;
 CC Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
 CC -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
 CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
 CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT (BY SIMILARITY).
 CC -1- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
 CC CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
 CC LINKS (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER COLLAGENS. STRONG, TO OTHER CUTICLE
 CC COLLAGENS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U08441; AAB37656.1;
 DR WormPep: F27C1.8; CE09720.
 DR InterPro: IPR000087; -
 DR Pfam: PF01484; Col-cuticle_N; 1.
 DR Pfam: PF01391; Collagen; 2.
 KW Hypothetical protein; Cuticle; Connective tissue; Repeat;
 KM Multigene family; Collagen.
 FT DOMAIN 94 126
 FT 143 270
 SO SEQUENCE 284 AA; 28216 MW; 8BED98AED554B454 CRC64;

Query Match
 Best Local Similarity 78.6%; Score 33; DB 1; Length 284;
 Matches 6; Conservative 0; Pred. No. 46; Mismatches 1; Indels 0; Gaps 0;

QY 2 VGVLRPG 8
 Db 211 VGVLRPG 217

RESULT 6
 YNCR_ECOLI
 ID YNCR_ECOLI STANDARD; PRT; 353 AA.
 AC F76113; P78255;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE PUTATIVE NADP-DEPENDENT OXIDOREDUCTASE IN TEHB-RHSE INTERGENIC REGION
 DE (EC 1.-.-.-).
 GN YNCR.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 RN NCBI_TaxID=562;
 RA
 RC STRAIN-K12 / MG1655;
 RC MEDLINE-97251357; PubMed-9097039;
 RX Blattner F.R., Plunkett G. III, Bloch C.A., Berra N.T., Burland V.,
 RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT The complete genome sequence of Escherichia coli K-12.;
 RL Science 277:1453-1474(1997).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RA MEDLINE-97251357; PubMed-9097039;
 RA Alpha H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
 RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
 RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubraman S.,
 RA Tagami H., Takeda Y., Takemoto K., Takeuchi Y., Wada C.,
 RA Yamamoto Y., Horuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 CC -1- SIMILARITY: BELONGS TO THE LABD FAMILY OF NADP-DEPENDENT
 CC OXIDOREDUCTASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE000242; AAC74531.1; ALT_INIT.
 DR EMBL: D90784; BAA15081.1;
 DR EMBL: D90785; BAA15084.1;
 DR Ecogene: EG13772; yncB.
 KW Hypothetical protein; Oxidoreductase; NADP.
 SO SEQUENCE 353 AA; 38495 MW; 120BEB0784311B55 CRC64;

Query Match
 Best Local Similarity 78.6%; Score 33; DB 1; Length 353;
 Matches 6; Conservative 1; Pred. No. 57; Mismatches 1; Indels 0; Gaps 0;

QY 1 VGVLRPG 8

DB 135 LGVLCMPG 142

```

RESULT 7
COLQ_TORMA
ID COLQ_TORMA STANDARD: PRT: 471 AA.
AC 003637:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE ACETYLCHOLINESTERASE COLLAGENIC TAIL PEPTIDE PRECURSOR (ACHE Q
SUBUNIT).
OS Torpedo marmorata (Marbled electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Hypnosqualia; Pristiogaster; Batoidae;
OC Torpediniformes; Torpedinidae; Torpedinidae; Torpedo.
OX NCBI_TaxID=7788;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-Electric Organ;
RA MEDLINE=92371432; PubMed=1840520;
RA Krejci E., Coussen F., Duval N., Chatelet J.-M., Legay C., Puype M.,
RA Vanderkerckhove J., Carlaud J., Bon S., Massoulie J.;
RT "Primary structure of a collagenic tail peptide of Torpedo
RT acetylcholinesterase: co-expression with catalytic subunit induces
RT the production of collagen-tailed forms in transfected cells.";
RL EMBO J. 10:11285-1293(1991).
RN [2]
RP SUBUNITS INTERACTIONS.
RC TISSUE-Electric Organ;
RA MEDLINE=92371432; PubMed=1840520;
RA Duval N., Krejci E., Grassi J., Coussen F., Massoulie J., Bon S.;
RT "Molecular architecture of acetylcholinesterase collagen-tailed
RT forms: construction of a glycolipid-tailed tetramer.";
RL EMBO J. 11:3255-3261(1992).
RN [3]
RP IDENTIFICATION OF PRO-RICH ATTACHMENT DOMAIN.
RA MEDLINE=97160616; PubMed=9006950;
RA Bon S., Coussen F., Massoulie J.;
RT "Quaternary associations of acetylcholinesterase. II. The polyproline
RT attachment domain of the collagen tail.";
RL J. Biol. Chem. 272:3016-3021(1997).
RN [4]
RP 3D-STRUCTURE MODELING OF COLLAGEN-LIKE DOMAIN.
RA MEDLINE=20146083; PubMed=10679527;
RA Deprez P., Inestrosa N.C.;
RT "Molecular modeling of the collagen-like tail of asymmetric
RT acetylcholinesterase.";
RL Protein Eng. 13:27-34(2000).
CC -1- FUNCTION: THE CATALYTIC SUBUNITS OF ASYMMETRIC ACHE TO THE
CC SYNAPTIC BASAL LAMINA.
CC -1- SUBUNIT: THE ASYMMETRIC FORM OF ACHE IS A DISULFIDE-BONDED
CC OLIGOMER COMPOSED OF A COLLAGENIC SUBUNIT (O) AND A VARIABLE
CC NUMBER OF ASYMMETRIC (T) CATALYTIC SUBUNITS. THE N-TERMINAL OF THE
CC COLLAGENIC SUBUNIT (O) ASSOCIATES WITH THE C-TERMINAL OF THE
CC CATALYTIC SUBUNIT (T).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ELECTRIC ORGANS BUT NOT IN
CC MUSCLE.
CC -1- DOMAIN: THE PROLINE-RICH ATTACHMENT DOMAIN (PRAD) BINDS THE ACHE
CC CATALYTIC SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE COLQ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X59359; CAA42009.1; -.
CC PIR: S15035; S15035.

```

```

DR InterPro: IPR000087; -.
DR Pfam: PF01391; Collagen. 2.
DR Signal: Synapse; Neurotransmitter degradation; Repeat; Collagen.
FT SIGNAL 1 30
FT CHAIN 31 471
FT DOMAIN 70 86
FT DOMAIN 118 282
FT DOMAIN 293 307
FT DOMAIN 75 84
FT DOMAIN 388 443
FT REPEAT 388 413
FT REPEAT 420 443
FT DISULFID 70 70
FT DISULFID 71 71
FT DISULFID 111 111
FT DISULFID 115 115
FT DISULFID 307 307
FT DISULFID 309 309
SQ SEQUENCE 471 AA; 49545 MW; B0137A393758D539 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 471;
Best Local Similarity 62.5%; Pred. No. 75;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVLRPG 8
Db 141 IGEIGRPG 148

RESULT 8
BISC_RHOSH
ID BISC_RHOSH STANDARD: PRT: 744 AA.
AC P54934;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BIOTIN SULFOXYDE REDUCTASE (EC 1.-.-) (BDS REDUCTASE) (BSO
DE REDUCTASE).
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=F. SP. DENTRIFICANS IL106;
RX MEDLINE=95251380; PubMed=7733660;
RA Pollock V.V., Barber M.J.;
RT "Molecular cloning and expression of biotin sulfoxide reductase from
RT Rhodobacter sphaeroides forma sp. dentrificans.";
RL Arch. Biochem. Biophys. 318:322-332(1995).
CC -1- FUNCTION: THIS ENZYME MAY SERVE AS A SCAVENGER, ALLOWING THE CELL
CC TO UTILIZE BIOTIN SULFOXYDE AS A BIOTIN SOURCE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: REDUCES A SPONTANEOUS OXIDATION PRODUCT OF
CC BIOTIN, D-BIOTIN D-SULFOXYDE (BSO OR BDS), BACK TO BIOTIN.
CC -1- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN).
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
CC OXIDOREDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: U08189; AAA74739.1; -.
CC HSSP: Q57366; ICXT.
DR InterPro: IPR001467; -.
DR Pfam: PF01568; Molybdop binding; 1.
DR Pfam: PF00384; molybdopterin; 1.

```

DR PROSITE: PS00551; MOLYBDOPTERIN_PROK_1; FALSE_NEG.
 DR PROSITE: PS00490; MOLYBDOPTERIN_PROK_2; 1.
 DR PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
 KW Olfactomeducase; Molybdenum.
 SQ SEQUENCE 744 AA; 80266 MW; 731A933E95358A19 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 744;
 Best Local Similarity 62.5%; Pred. NO. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVLRPG 8
 Db 320 IGOIGRPG 327

RESULT 9
 CA14_MOUSE STANDARD; PRT; 1669 AA.

AC P02463;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE COLLAGEN ALPHA 1(IV) CHAIN PRECURSOR.
 GN COL4A1
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN 1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89197932; PubMed-2703490;
 RA Muthukumaran G., Blumberg B., Kurkinen M.;
 RT "The complete primary structure for the alpha 1-chain of mouse
 RT collagen IV. Differential evolution of collagen IV domains.";
 RL J. Biol. Chem. 264:6310-6317(1989).
 RP SEQUENCE OF 1-1154 FROM N.A.
 RA MEDLINE=88112221; PubMed-3338568;
 RA Wood L., Theriault N., Vogeli G.;
 RT "cDNA clones completing the nucleotide and derived amino acid
 RT sequence of the alpha 1 chain of basement membrane (type IV) collagen
 RT from mouse.";
 RL FEBS Lett. 227:5-8(1988).
 RP SEQUENCE OF 1149-1424 FROM N.A.
 RA MEDLINE=86301886; PubMed-3755692;
 RA Nath P., Laurent M., Horn E., Sobel M.E., Zon G., Vogeli G.;
 RT "Isolation of an alpha 1 type-IV collagen cDNA clone using a
 RT synthetic oligodeoxynucleotide.";
 RL Gene 43:301-304(1986).
 RP SEQUENCE OF 1276-1669 FROM N.A.
 RA MEDLINE=85127033; PubMed-2578961;
 RA Oberbauer I., Laurent M., Schwarz U., Sakurai Y., Yamada Y.,
 RA Vogeli G., Voss T., Siebold B., Glaville R.W., Kuhn K.;
 RT "Amino acid sequence of the non-collagenous globular domain (NC1) of
 RT the alpha 1(IV) chain of basement membrane collagen as derived from
 RT complementary DNA.";
 RL Eur. J. Biochem. 147:217-224(1985).
 RP SEQUENCE OF 1441-1669 FROM N.A.
 RA MEDLINE=87250460; PubMed-3597383;
 RA Kurkinen M., Condon M.R., Blumberg B., Barlow D., Quinones S.,
 RA Saus J., Pihlajaniemi T.;
 RT "Extensive homology between the carboxyl-terminal peptides of mouse
 RT alpha 1(IV) and alpha 2(IV) collagen.";
 RL J. Biol. Chem. 263:8496-8499(1987).
 RP PARTIAL SEQUENCE FROM N.A.
 RA MEDLINE=86196099; PubMed-3009468;
 RA Sakurai Y., Sullivan M., Yamada Y.;
 RT "Alpha 1 type IV collagen gene evolved differently from fibrillar

RT collagen genes.";
 RL J. Biol. Chem. 261:6654-6657(1986).
 RP SEQUENCE OF 1-28 FROM N.A.
 RA MEDLINE=89066738; PubMed-3198626;
 RA Kaytes P., Wood L., Theriault N., Kurkinen M., Vogeli G.;
 RT "Head-to-head arrangement of murine type IV collagen genes.";
 RL J. Biol. Chem. 263:19274-19277(1988).
 RP SEQUENCE OF 1-28 FROM N.A.
 RA MEDLINE=89071759; PubMed-3200851;
 RA Burbello P.D., Martin G.R., Yamada Y.;
 RT "Alpha 1(IV) and alpha 2(IV) collagen genes are regulated by a
 RT bidirectional promoter and a shared enhancer.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:9679-9682(1988).
 RP SEQUENCE OF 1-129 FROM N.A.
 RA MEDLINE=88243724; PubMed-3379041;
 RA Killen P.D., Burbello P., Sakurai Y., Yamada Y.;
 RT "Structure of the amino-terminal portion of the murine alpha 1(IV)
 RT collagen chain and the corresponding region of the gene.";
 RL J. Biol. Chem. 263:8706-8709(1988).
 RP FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
 CC GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
 CC MESHWORK TOGETHER WITH LAMININS, PROTEOLYCAN AND ENACTIN/
 CC NIDOGEN.
 CC -1- SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
 CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
 CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
 CC -1- DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 CC DOMAIN (NC1) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 CC G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CC CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
 CC TRIPLE-HELICAL 7S DOMAIN.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
 CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 CC THESE, LOCATED IN THE NC1 DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 CC IV COLLAGENS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: J03758; AAA37439.1; -
 DR EMBL: M23333; AAA51625.1; -
 DR EMBL: J04694; AAA50292.1; -
 DR EMBL: X06777; CAA29946.1; -
 DR EMBL: X02201; CAA26132.1; -
 DR EMBL: M15832; AAA37340.1; -
 DR EMBL: M14042; AAA37342.1; -
 DR EMBL: M12879; AAA37343.1; -
 DR EMBL: M3024; -; NOT_ANNOTATED_CDS.
 DR EMBL: M13025; -; NOT_ANNOTATED_CDS.
 DR EMBL: M13026; AAA37344.1; -
 DR EMBL: M13027; AAA37345.1; -
 DR EMBL: M13043; AAA37346.1; -
 DR EMBL: J04448; AAA37437.1; -
 DR PIR: A33525; CGMS4B.
 DR MGD: MGI:88454; Col4a1.
 DR InterPro: IPR000087; -
 DR InterPro: IPR001442; -
 DR Pfam: PF01413; C4; 2.
 DR Pfam: PF01391; Collagen; 21.
 KW Extracellular matrix; Connective tissue; Basement membrane;
 KW Repeat; Hydroxylation; Glycoprotein; Collagen; Signal.
 FT SIGNAL
 FT PROPEP
 28 172 AMINO-TERMINAL PROPEPTIDE (7S DOMAIN).
 27

```

FT CHAIN 173 1669 COLLAGEN ALPHA 1(IV) CHAIN.
FT DOMAIN 173 1440 TRIPLE-HELICAL REGION.
FT DOMAIN 1441 1669 NONHELICAL REGION (NC1).
FT DISULFID 1460 1551 OR 1548 (BY SIMILARITY).
FT DISULFID 1493 1548 OR 1551 (BY SIMILARITY).
FT DISULFID 1505 1511 BY SIMILARITY.
FT DISULFID 1570 1665 OR 1662 (BY SIMILARITY).
FT DISULFID 1604 1662 OR 1665 (BY SIMILARITY).
FT DISULFID 1616 1622 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 26 26 A -> P (IN REF. 2).
FT CONFLICT 186 186 S -> L (IN REF. 2).
FT CONFLICT 319 319 O -> S (IN REF. 2).
FT CONFLICT 369 369 Q -> L (IN REF. 2).
FT CONFLICT 403 403 L -> F (IN REF. 2).
FT CONFLICT 481 481 P -> L (IN REF. 2).
FT CONFLICT 493 493 Q -> H (IN REF. 2).
FT CONFLICT 712 712 S -> I (IN REF. 2).
FT CONFLICT 813 813 E -> Q (IN REF. 2).
FT CONFLICT 982 982 Q -> H (IN REF. 2).
FT CONFLICT 1397 1397 V -> S (IN REF. 3).
SQ SEQUENCE 1669 AA; 160680 MM; 42916B91E52058B9 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 1669;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVLRPG 8
DB 1057 IGIPGRG 1064

RESULT 10
ID CAL5_HUMAN STANDARD; PRT; 1838 AA.
AC P20908;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE COLLAGEN ALPHA 1(V) CHAIN PRECURSOR.
GN COL5A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
OX 11;
RN RP SEQUENCE FROM N.A., AND SEQUENCE OF 556-565.
RX MEDLINE=91302336; PubMed=2071595;
RA Takahara K., Seto Y., Okasawa K., Okamoto N., Noda A., Yaol Y.,
RA Kato I.;
RT "Complete primary structure of human collagen alpha 1 (V) chain.";
RL J. Biol. Chem. 266:13124-13129(1991).
RN 12;
RP SEQUENCE OF 621-822.
RC TISSUE=Chorioallantoic membrane;
RX MEDLINE=89227189; PubMed=2496661;
RA Seyer J.M., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of three
RT cyanogen bromide-derived peptides from human alpha 1(V) collagen
RT chain.";
RL Arch. Biochem. Biophys. 271:120-129(1989).
RN 13;
RP SEQUENCE OF 823-950, AND HEPARIN-BINDING.
RX MEDLINE=90366601; PubMed=2203476;
RA Yaol Y., Hashimoto K., Koitabashi H., Takahara K., Ito M., Kato I.;
RT "Primary structure of the heparin-binding site of type V collagen.";
RL Biochim. Biophys. Acta 1035:139-145(1990).
RN 14;
RP SEQUENCE OF 556-571.
RC TISSUE=Placenta;
RX MEDLINE=92239022; PubMed=1571108;
RA Mann K.;

```

```

RT "Isolation of the alpha 3-chain of human type V collagen and
RT characterization by partial sequencing.";
RL Biol. Chem. Hoppe-Seyler 373:69-75(1992).
RN 15;
RP SEQUENCE OF 565-576; 756-772; 1012-1029; 1219-1232 AND 1465-1477.
RC TISSUE=Chorioallantoic membrane;
RX MEDLINE=94237164; PubMed=8181482;
RA Bourdillon M.M., Rousseau J.C., Klemm J.P., Champplaud M.F.,
RA Boutillon M.M., Bernillon J., Wallach J.M., van der Rest M.;
RT "Diversity in the processing events at the N-terminus of type-V
RT collagen.";
RL Eur. J. Biochem. 221:987-995(1994).
RN 16;
RP VARIANT EDS1 SER-1639.
RX MEDLINE=97195540; PubMed=9042913;
RA de Paeppe A., Nuytink L., Hauser I., Anton-Lamprecht I.,
RA Naeyaert J.-M.;
RT "Mutations in the COL5A1 gene are causal in the Ehlers-Danlos
RT syndromes I and II.";
RL Am. J. Hum. Genet. 60:547-554(1997).
CC -1- FUNCTION: TYPE V COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC (FIBRILLAR FORMING COLLAGEN). IT IS A MINOR CONNECTIVE TISSUE
CC COMPONENT OF NEARLY UBIQUITOUS DISTRIBUTION. TYPE V COLLAGEN BINDS
CC TO DNA, HEPARAN SULFATE, THROMBOSPONDIN, HEPARIN, AND INSULIN.
CC -1- SUBUNIT: PRIMERS OF TWO ALPHA 1(V) AND ONE ALPHA 2(V) CHAINS IN
CC MOST TISSUES AND TRIMERS OF ONE ALPHA 1(V), ONE ALPHA 2(V), AND
CC ONE ALPHA 3(V) CHAINS IN PLACENTA.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- PTM: 40% OF TYROSINES IN THE PRO-ALPHA 1(V) CHAIN ARE SULFATED.
CC -1- DISEASE: DEFECTS IN COL5A1 ARE A CAUSE OF EHLERS-DANLOS SYNDROME,
CC TYPE I (EDS1), A DISEASE CHARACTERIZED BY LOOSE-JOINTEDNESS AND
CC FRAGILE, VELVET, STRETCHABLE, BRUISEABLE SKIN THAT HEALS WITH
CC PECULIAR 'CIGARETTE-PAPER' SCARS.
CC -1- SIMILARITY: HIGH, TO ALPHA 3(V) AND ALPHA 1(XI) CHAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90279; BAA14323.1; -
DR PIR: S03978; S03978.
DR PIR: S11303; S11303.
DR PIR: S16024; S16024.
DR MIM: 120215; -
DR MIM: 130000; -
DR MIM: 130010; -
DR InterPro: IPR000087; -
DR InterPro: IPR000885; -
DR Pfam: PF01410; COLF1.1.
DR Pfam: PF01391; Collagen; 18.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Collagen; Signal; Heparin-binding; Sulfation; Disease mutation.
FT CHAIN 1 37
FT SIGNAL 1 37
FT DOMAIN 38 1605
FT DOMAIN 38 443
FT DOMAIN 444 558
FT DOMAIN 559 1570
FT DOMAIN 1571 1605
FT PROPEP 1606 1838
FT MOD_RES 570 570
FT MOD_RES 576 576
FT MOD_RES 621 621
FT MOD_RES 627 627
FT MOD_RES 639 639
FT MOD_RES 642 642
FT MOD_RES 648 648
FT MOD_RES 654 654
FT MOD_RES 657 657
FT MOD_RES 657 657

```

```

FT MOD_RES 675 675 HYDROXYLATION.
FT MOD_RES 678 678 HYDROXYLATION.
FT MOD_RES 680 680 HYDROXYLATION.
FT MOD_RES 686 686 HYDROXYLATION.
FT MOD_RES 690 690 HYDROXYLATION.
FT MOD_RES 696 696 HYDROXYLATION.
FT MOD_RES 705 705 HYDROXYLATION.
FT MOD_RES 708 708 HYDROXYLATION.
FT MOD_RES 717 717 HYDROXYLATION.
FT MOD_RES 720 720 HYDROXYLATION.
FT MOD_RES 726 726 HYDROXYLATION.
FT MOD_RES 732 732 HYDROXYLATION.
FT MOD_RES 744 744 HYDROXYLATION.
FT MOD_RES 750 750 HYDROXYLATION.
FT MOD_RES 756 756 HYDROXYLATION.
FT MOD_RES 762 762 HYDROXYLATION.
FT MOD_RES 765 765 HYDROXYLATION.
FT MOD_RES 771 771 HYDROXYLATION.
FT MOD_RES 774 774 HYDROXYLATION.
FT MOD_RES 780 780 HYDROXYLATION.
FT MOD_RES 789 789 HYDROXYLATION.
FT MOD_RES 795 795 HYDROXYLATION.
FT MOD_RES 804 804 HYDROXYLATION.
FT MOD_RES 807 807 HYDROXYLATION.
FT MOD_RES 810 810 HYDROXYLATION.
FT MOD_RES 816 816 HYDROXYLATION.
FT MOD_RES 819 819 HYDROXYLATION.
FT MOD_RES 834 834 HYDROXYLATION.
FT MOD_RES 846 846 HYDROXYLATION.
FT MOD_RES 861 861 HYDROXYLATION.
FT MOD_RES 864 864 HYDROXYLATION.
FT MOD_RES 870 870 HYDROXYLATION.
FT MOD_RES 873 873 HYDROXYLATION.
FT MOD_RES 876 876 HYDROXYLATION.
FT MOD_RES 882 882 HYDROXYLATION.
FT MOD_RES 888 888 HYDROXYLATION.
FT MOD_RES 891 891 HYDROXYLATION.
FT MOD_RES 897 897 HYDROXYLATION.
FT MOD_RES 903 903 HYDROXYLATION.
FT MOD_RES 906 906 HYDROXYLATION.
FT MOD_RES 930 930 HYDROXYLATION.
FT MOD_RES 945 945 HYDROXYLATION.
FT MOD_RES 1017 1017 HYDROXYLATION.
FT MOD_RES 1020 1020 HYDROXYLATION.
FT MOD_RES 1023 1023 HYDROXYLATION.
FT MOD_RES 1029 1029 HYDROXYLATION.
FT MOD_RES 1221 1221 HYDROXYLATION.
FT MOD_RES 1224 1224 HYDROXYLATION.
FT MOD_RES 1467 1467 HYDROXYLATION.
FT MOD_RES 1470 1470 HYDROXYLATION.
FT VARIANT 1639 1639
FT CONFLICT 641 641 /FTID=VAR.001808.
FT CONFLICT 650 650 E -> G (IN REF. 2).
FT CONFLICT 663 663 P -> L (IN REF. 2).
FT CONFLICT 668 668 R -> E (IN REF. 2).
FT CONFLICT 677 677 K -> Q (IN REF. 2).
FT CONFLICT 684 684 L -> P (IN REF. 2).
FT CONFLICT 692 692 L -> P (IN REF. 2).
FT CONFLICT 699 699 PGPSPVT -> VTGSPGAP (IN REF. 2).
FT CONFLICT 727 727 G -> Q (IN REF. 2).
FT CONFLICT 741 741 P -> L (IN REF. 2).
FT CONFLICT 747 747 L -> Q (IN REF. 2).
FT CONFLICT 753 753 P -> A (IN REF. 2).
FT CONFLICT 759 759 D -> N (IN REF. 2).
FT CONFLICT 776 776 GQ -> QK (IN REF. 2).
FT CONFLICT 849 849 GGPSPDP -> IGPPGPR (IN REF. 3).
FT CONFLICT 894 894 N -> D (IN REF. 3).
SO SEQUENCE 1838 AA; 183616 MM; 7D58239C0D7BDAE CRC64;

```

Query Match 78.6%; Score 33; DB 1; Length 1838;
 Best Local Similarity 62.5%; Pred. No. 2.7e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 VGVLRPG 8
Db 1101 IGIPCRPG 1108

RESULT 11
POLN.SINDV
AC P0317; STANDARD; PRT; 2512 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE NONSTRUCTURAL POLYPEPTIDE SEQUENCE OF THE GENOMIC RNA OF SINDHIS VIRUS.
DE NSP1 TO NSP4.
OS Sindhis virus (strain HRSP).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OX Alphavirus.
RN NCBL_TaxID=11034;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=84148439; PubMed=6322438;
RA Strauss E.G., Rice C.M., Strauss J.H.;
RT "Complete nucleotide sequence of the genomic RNA of Sindhis virus."
RL Virology 133:92-110(1984).
RN [2]
RN SEQUENCE OF 1-54 FROM N.A.
RX MEDLINE=83268700; PubMed=6308269;
RA Ou J.H., Strauss E.G., Strauss J.H.;
RT "The 5'-terminal sequences of the genomic RNAs of several
  J. Mol. Biol. 168:1-15(1983).
RN [3]
RN SEQUENCE OF 1429-2512 FROM N.A.
RX MEDLINE=83039446; PubMed=6291034;
RA Ou J.H., Rice C.M., Dalgarno L., Strauss E.G., Strauss J.H.;
RT "Sequence studies of several alphavirus genomic RNAs in the region
  containing the start of the subgenomic RNA."
RL Proc. Natl. Acad. Sci. U.S.A. 79:5235-5239(1982).
CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS.
CC BETWEEN THE CODONS FOR 1896-TYR AND 1897-LEU.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: J02363; AAA56975.1; ALT_FRAME.
CC PIR: A03917; MNWVS.
CC MEROPS: C09.001; -.
CC InterPro: IPR000606; -.
CC InterPro: IPR001788; -.
CC InterPro: IPR002589; -.
CC InterPro: IPR002620; -.
CC pfam: PF01661; DUF27.1.
CC pfam: PF01707; Peptidase_C9; 1.
CC pfam: PF00978; RNA_dep_RNApol2; 2.
CC pfam: PF01443; Viral_helicase1; 1.
CC PolyProtein; Nonstructural protein; RNA-binding; Helicase.
FT CHAIN 1 540 NONSTRUCTURAL PROTEIN NSP1.
FT CHAIN 541 1347 NONSTRUCTURAL PROTEIN NSP2.

```

FT CHAIN 1348 1896 NONSTRUCTURAL PROTEIN NSP3.
 FT CHAIN 1897 2512 NONSTRUCTURAL PROTEIN NSP4.
 SQ SEQUENCE 2512 AA: 279546 MW: F3656FCBB495726 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 2512;
 Best Local Similarity 62.5%; Pred No. 3.6e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VGVLCRPG 8
 :||:|
 Db 722 IGVIGRPG 729

RESULT 12
 POLN_SINDO STANDARD: PRT; 2514 AA.
 AC P27283;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE NONSTRUCTURAL POLYPEPTIDE (P270) [CONTAINS: NONSTRUCTURAL PROTEINS
 DE NSP3 TO NSP4].
 OS Sindbis virus (subtype Ockelbo / strain Edsbyn 82-5).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OX NCBI_TaxID=31699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9120725; PubMed=1673813;
 RA Shtrako Y., Niklason B., Dalrymple J.M., Strauss E.G., Strauss J.H.;
 RT "Structure of the Ockelbo virus genome and its relationship to other
 RT Sindbis viruses.";
 RL Virology 182:753-764(1991).
 CC -1- FUNCTION: NSP2 MAY BE INVOLVED IN RNA BINDING DURING REPLICATION.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UGA OCCURS
 CC BETWEEN THE CODONS FOR 1898-TYR AND 1899-LEU.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M69205; AAA96972.1; ALT_FRAME.
 DR PIR: A39991; MNAV82.
 DR MEROPS: C09.001; -;
 DR InterPro: IPR000606; -;
 DR InterPro: IPR001788; -;
 DR InterPro: IPR002589; -;
 DR InterPro: IPR002620; -;
 DR Pfam: PF01661; DUF27.1.
 DR Pfam: PF01707; Peptidase_C9.1.
 DR Pfam: PF00978; RNA_dep_RNApol2.2.
 DR Pfam: PF01443; Viral_helicase1.1.
 KW Polypeptide; Nonstructural protein; RNA-binding; Helicase.
 FT CHAIN 1 540
 FT CHAIN 541 1347 NONSTRUCTURAL PROTEIN NSP1.
 FT CHAIN 1348 1898 NONSTRUCTURAL PROTEIN NSP2.
 FT CHAIN 1899 2514 NONSTRUCTURAL PROTEIN NSP4.
 FT CHAIN 2514 AA: 279642 MW: 2F38CE32ACF5EDD CRC64;
 SQ SEQUENCE

Query Match 78.6%; Score 33; DB 1; Length 2514;
 Best Local Similarity 62.5%; Pred No. 3.6e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VGVLCRPG 8
 :||:|
 Db 722 IGVIGRPG 729

RESULT 13
 CA17_HUMAN STANDARD: PRT; 2944 AA.
 AC 002388; 014054; 016507;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE COLLAGEN ALPHA 1(VII) CHAIN PRECURSOR (LONG-CHAIN COLLAGEN) (LC
 DE COLLAGEN).
 DE COL7A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94327588; PubMed=8051117;
 RA Christiano A.M., Greenspan D.S., Lee S., Uitto J.;
 RT "Cloning of human type VII collagen. Complete primary sequence of the
 RT alpha 1(VII) chain and identification of intragenic polymorphisms.";
 RL J. Biol. Chem. 269:20256-20262(1994).
 RN [2]
 RP SEQUENCE OF 128-1493 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93338437; PubMed=1307247;
 RA Christiano A.M., Rosenbaum L.M., Chung-Honet L.C., Parente M.G.,
 RA Woodley D.T., Pan T.C., Zhang R.Z., Chu M.L., Burgess R.E.,
 RA Uitto J.;
 RT "The large non-collagenous domain (NC-1) of type VII collagen is
 RT amino-terminal and chimeric. Homology to cartilage matrix protein,
 RT the type III domains of fibronectin and the A domains of von
 RT Willebrand factor.";
 RL Hum. Mol. Genet. 1:475-481(1992).
 RN [3]
 RP SEQUENCE OF 815-1439 FROM N.A.
 RX MEDLINE=91334380; PubMed=1871109;
 RA Parente M.G., Chung L.C., Rymaszewski J., Woodley D.T., Wynn K.W.,
 RA Bauer E.A., Mattei M.-G., Chu M.-L., Uitto J.;
 RT "Human type VII collagen: cDNA cloning and chromosomal mapping of the
 RT gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:6931-6935(1991).
 RN [4]
 RP SEQUENCE OF 369-1255 FROM N.A.
 RX MEDLINE=93107742; PubMed=1469284;
 RA Gannon W.R., Abernethy M.L., Padilla K.M., Prisyah P.S.,
 RA Cook M.E., Wright J., Brigaman R.A., Hunt S.W. II;
 RT "Noncollagenous (NC1) domain of collagen VII resembles multidomain
 RT adhesion proteins involved in tissue-specific organization of
 RT extracellular matrix.";
 RL J. Invest. Dermatol. 99:691-696(1992).
 RN [5]
 RP SEQUENCE OF 340-675 FROM N.A.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=92231902; PubMed=1567409;
 RA Tanaka T., Takahashi K., Furukawa F., Imamura S.;
 RT "Molecular cloning and characterization of type VII collagen cDNA.";
 RL Biochem. Biophys. Res. Commun. 183:958-963(1992).
 RN [6]
 RP SEQUENCE OF 2395-2944 FROM N.A.
 RX MEDLINE=93271985; PubMed=849916;
 RA Greenspan D.S.;
 RT "The carboxyl-terminal half of type VII collagen, including the non-
 RT collagenous NC-2 domain and intron/exon organization of the
 RT corresponding region of the COL7A1 gene.";
 RL Hum. Mol. Genet. 2:273-278(1993).
 RN [7]
 RP SEQUENCE OF 1-87 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94375010; PubMed=8088784;
 RA Christiano A.M., Hoffman G.G., Chung-Honet L.C., Lee S., Cheng W.,
 RA Uitto J., Greenspan D.S.;
 RT "Structural organization of the human type VII collagen gene (COL7A1),

RT composed of more exons than any previously characterized gene.";
RL Genomics 21:169-179(1994).
RN [8].
RP REVIEW ON DEB VARIANTS.
RX MEDLINE=98041696; PubMed=9375848.
RA Jaervikallio A., Pulkkinen L., Uitto J.;
RT "Molecular basis of dystrophic epidermolysis bullosa: mutations in
RL the type VII collagen gene (COL7A1).";
RN Hum. Mutat. 10:358-347(1997).
RN [9].
RP VARIANT RDEB LYS-2798.
RX MEDLINE=93291877; PubMed=8513326;
RA Christlano A.M., Greenspan D.S., Hoffman G.G., Zhang X., Tamai Y.,
RA Lin A.N., Dietz H.C., Hovnanian A., Uitto J.;
RT "A missense mutation in type VII collagen in two affected sibs
RN with recessive dystrophic epidermolysis bullosa.";
RL Nat. Genet. 4:62-66(1993).
RP VARIANT DDEB SER-2040.
RX MEDLINE=94224777; PubMed=8170945;
RA Christlano A.M., Ryyanen M., Uitto J.;
RT "Dominant dystrophic epidermolysis bullosa: identification of a
RN glycine-to-serine substitution in the triple-helical domain of type VII
RL collagen.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3549-3553(1994).
RP [11].
RP VARIANT PER-DDEB CYS-2633.
RX MEDLINE=96081220; PubMed=8541942;
RA Christlano A.M., Lee J.Y.-Y., Chen W.J., Laforgia S., Uitto J.;
RT "Pretibial epidermolysis bullosa: genetic linkage to COL7A1 and
RN identification of a glycine-to-cysteine substitution in the triple-
RL helical domain of type VII collagen.";
RN Hum. Mol. Genet. 4:1579-1583(1995).
RN [12].
RP VARIANT DDEB ARG-2043.
RX MEDLINE=95164385; PubMed=7861014;
RA Christlano A.M., Morricone A., Paradisi M., Angelo C., Mazzanti C.,
RA Chavaleret R., Uitto J.;
RT "A glycine-to-arginine substitution in the triple-helical domain of
RN type VII collagen in a family with dominant dystrophic epidermolysis
RL bullosa.";
RL J. Invest. Dermatol. 104:438-440(1995).
RN [13].
RP VARIANTS RDEB AND DDEB.
RX MEDLINE=96220218; PubMed=8644729;
RA Christlano A.M., McGrath J.A., Yan K.C., Uitto J.;
RT "Glycine substitutions in the triple-helical region of type VII
RN collagen result in a spectrum of dystrophic epidermolysis bullosa
RL phenotypes and patterns of inheritance.";
RN Am. J. Hum. Genet. 58:671-681(1996).
RN [14].
RP VARIANT RDEB ARG-2575.
RX MEDLINE=96154068; PubMed=8592061;
RA Shimizu H., McGrath J.A., Christlano A.M., Nishikawa T., Uitto J.;
RT "Molecular basis of recessive dystrophic epidermolysis bullosa:
RN genotype/phenotype correlation in a case of moderate clinical
RL severity.";
RL J. Invest. Dermatol. 106:119-124(1996).
RN [15].
RP VARIANT RDEB ARG-1782.
RX MEDLINE=96183562; PubMed=8618018;
RA Christlano A.M., McGrath J.A., Uitto J.;
RT "Influence of the second COL7A1 mutation in determining the
RN phenotypic severity of recessive dystrophic epidermolysis bullosa.";
RL J. Invest. Dermatol. 106:766-770(1996).
RN [16].
RP VARIANT RDEB ASP-2073.
RX MEDLINE=96310789; PubMed=8757758;
RA Dunnill M.G.S., McGrath J.A., Richards A.J., Christlano A.M.,
RA Uitto J., Pope F.M., Eady R.A.J.;
RT "Clinicopathological correlations of compound heterozygous COL7A1
RN mutations in recessive dystrophic epidermolysis bullosa.";
RL J. Invest. Dermatol. 107:171-177(1996).

RP VARIANTS RDEB W-1982; G-2008; A-2025; E-2049; G-2063; W-2063 & R-2575.
 RX MEDLINE=97465605; PubMed=9326325.
 RA Hovenian A., Rochat A., Bodemer C., Petit E., Rivers C.A., Prost C.,
 RA Fraitag S., Cristiano A.M., Uitto J., Lathrop M., Barrardon Y.,
 RA de Prost Y.:
 "Characterization of 18 new mutations in COL7A1 in recessive
 RT dystrophic epidermolysis bullosa provides evidence for distinct
 RT molecular mechanisms underlying defective anchoring fibril
 RT formation.";
 RL Am. J. Hum. Genet. 61:599-610(1997).
 CC -1- FUNCTION: STRATIFIED SQUAMOUS EPITHELIAL BASEMENT MEMBRANE PROTEIN
 CC THAT FORM ANCHORING FIBRILS WHICH MAY CONTRIBUTE TO EPITHELIAL
 CC BASEMENT MEMBRANE ORGANIZATION AND ADHERENCE BY INTERACTING WITH
 CC EXTRACELLULAR MATRIX (ECM) PROTEINS SUCH AS TYPE IV COLLAGEN.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- PMV: PROLINS AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- DISEASE: DEFECT IN COL7A1 IS A CAUSE OF RECESSIVE DYSTROPHIC
 CC CHARACTERIZED BY TISSUE SEPARATION WITHIN THE DERMAL-EPIDERMAL
 CC BASEMENT MEMBRANE ZONE. THE DYSTROPHIC TYPE, THE SCARRING FORM
 CC MANIFESTS THE CLEAVAGE WITHIN THE PAPILLARY DERMS BELOW THE BASAL
 CC LAMINA AT THE LEVEL OF THE ANCHORING FIBRILS. BOTH AUTOSOMAL
 CC DOMINANT AND RECESSIVE INHERITANCE PATTERNS HAVE BEEN RECOGNIZED.
 CC -1- DISEASE: EPIDERMOLYSIS BULLOSA ACQUISITA (EBA) IS AN AUTOIMMUNE-
 CC ACQUIRED BLISTERING SKIN DISEASE RESULTING FROM AUTOANTIBODIES TO
 CC TYPE VII COLLAGEN.
 CC -1- DISEASE: DEFECT IN COL7A1 IS A CAUSE OF DYSTROPHIC EPIDERMOLYSIS
 CC BULLOSA (DEB). DEB IS A BLISTERING SKIN DISEASES WITH EITHER AN
 CC AUTOSOMAL DOMINANT (DDEB) OR AUTOSOMAL RECESSIVE (RDEB)
 CC INHERITANCE. ULTRASTRUCTURALLY, DEB IS CHARACTERIZED BY TISSUE
 CC SEPARATION WHICH OCCURS BELOW THE DERMAL-EPIDERMAL BASEMENT
 CC MEMBRANE AT THE LEVEL OF THE ANCHORING FIBRILS. THE MOST SEVERE
 CC FORM OF RDEB IS THE HALLOPEAU-STEMENS TYPE (HS-RDEB), A MILDER
 CC AND THE LOCALIZED TYPES (M-RDEB), AND THERE ARE STILL THE INVERSE,
 CC MUTILATING SCARRING, WITH JOINT CONTRACTURES, CORNEAL EROSIONS,
 CC ESOPHAGUS STENOSIS, AND PROPENSITY TO FORMATION OF CUTANEOUS
 CC SQUAMOUS CELL CARCINOMAS LEADING TO PREMATURE DEMISE OF THE
 CC WITH LIMITED SCARRING. M-RDEB SHOWS LIFELONG BLISTERING TENDENCY,
 CC MANIFESTATIONS. HOWEVER THERE IS A CONTINUUM OF SEVERITY BETWEEN
 CC HS AND M-RDEB. THERE ARE SEVERAL TYPES OF DDEB: THE COCKAYNE-
 CC TOURNAI TYPE (CT-DDEB), THE PASINI TYPE (P-DDEB) WHICH IS SEVERE,
 CC BARTH TYPE (B-DDEB) AND PRETIBIAL EPIDERMOLYSIS BULLOSA
 CC (PBP-DDEB).
 CC -1- SIMILARITY: THE NC-1 DOMAIN OF TYPE VII COLLAGEN HAS HOMOLOGY
 CC TO FIBRONECTIN TYPE III DOMAINS, VON WILLEBRAND FACTOR A
 CC DOMAINS AND CARTILAGE MATRIX PROTEIN.
 CC -1- SIMILARITY: CONTAINS 1 PROTEASE INHIBITOR DOMAIN BELONGING TO THE
 CC BPTI/KUNITZ FAMILY OF INHIBITORS.
 CC -1- SIMILARITY: CONTAINS 9 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- CAUTION: REP 5 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS
 CC 476 TO 523 DUE TO A FRAMESHIFT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC

Query Match 78.6%; Score 33; DB 1; Length 2944;
 Best Local Similarity 85.7%; Pred. No. 4.1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 2 GYGRCR 8
 Db 1510 GYAGRCR 1516

RESULT 14
 ID RL11_PIG
 AC G29205; STANDARD: PRT; 165 AA.
 JT 15-DEC-1998 (Rel. 37, Created)

```

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S RIBOSOMAL PROTEIN L11 (FRAGMENT).
GN RPL11.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxId=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Wintere A.K., Fredholm M., Davies M.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RL library: analysis of 839 clones.";
RM Mamm. Genome 7:509-517(1996).
CC -1- FUNCTION: PROBABLY ASSOCIATES WITH 5S RNA.
CC -1- SIMILARITY: BELONGS TO THE LSP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: F14532; CAA23111.1; -
DR InterPro: IPR002132; -
DR Pfam: PF00281; Ribosomal_L5; 1.
DR Pfam: PF00673; Ribosomal_L5.C; 1.
DR PROSITE: PS00358; RIBOSOMAL_L5; 1.
KW Ribosomal protein; rRNA-binding.
FT INIT_MET 0
FT NON_TER 165
FT SEQUENCE 165 AA; 18766 MW; F269DDEDFE551F0 CRC64;
SQ
Query Match 76.2%; Score 32; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 VLGRPG 8
DB 1332 VLGRPG 137

```

```

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 16-51 FROM N.A.
RC SPECIES=Human;
RX MEDLINE=98248690; PubMed=9582194;
RA Kenmochi N., Kawaguchi T., Rozen S., Davis E., Goodman N.,
RA Hudson T.J., Tanaka T., Page D.C.;
RT "A map of 75 human ribosomal protein genes.";
RL Genome Res. 8:509-523(1998).
RN [4]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC SPECIES=Rat; STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE=92287119; PubMed=1599472;
RA Chan Y.-L., Olivera J., Paz V., Wool I.G.;
RA Biochem. Biophys. Res. Commun. 185:356-362(1992).
RL The primary structure of rat ribosomal protein L11."
CC -1- FUNCTION: PROBABLY ASSOCIATES WITH 5S RNA.
CC -1- SIMILARITY: BELONGS TO THE LSP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X79234; CAA55816.1; -
DR EMBL: L05092; AAC15856.1; -
DR EMBL: AB007171; BAA25831.1; -
DR EMBL: X62146; CAA44072.1; -
DR PIR: S17351; P5R711.
DR PIR: S45049; S45049.
DR PIR: J70606; J70606.
DR MIM: 604175; -
DR InterPro: IPR002132; -
DR Pfam: PF00281; Ribosomal_L5; 1.
DR Pfam: PF00673; Ribosomal_L5.C; 1.
DR PROSITE: PS00358; RIBOSOMAL_L5; 1.
KW Ribosomal protein; rRNA-binding.
FT INIT_MET 0
FT CONFICT 30
FT CONFICT 30 D -> G (IN REF. 1).
FT CONFICT 72 T -> A (IN REF. 1).
FT CONFICT 91 Y -> L (IN REF. 1).
FT CONFICT 91
FT CONFICT 117 K -> E (IN REF. 1).
FT SEQUENCE 177 AA; 20121 MW; 5BDBCBDE20F2D507 CRC64;
SQ
Query Match 76.2%; Score 32; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 VLGRPG 8
DB 1332 VLGRPG 137

```

Search completed: June 13, 2001, 14:30:38
Job time: 528 sec

Wed Jun 13 15:00:45 2001

pct-us01-05825a-29.rsp

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:29:45 ; Search time 150.43 Seconds
(without alignments)
6.233 Million cell updates/sec

Title: PCT-US01-05825A-29
Perfect score: 42
Sequence: 1 VGVLCRPG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP-archaea: *
2: SP-bacteria: *
3: SP-fungi: *
4: SP-human: *
5: SP-invertebrate: *
6: SP-mammal: *
7: SP-mhc: *
8: SP-organelle: *
9: SP-phage: *
10: SP-plant: *
11: SP-rodent: *
12: SP-unclassified: *
13: SP-vertebrate: *
14: SP-virus: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	90.5	812	5	006452 ephydattia m
2	37	88.1	189	5	097409 anopheles g
3	37	88.1	472	2	033954 streptomyce
4	36	85.7	577	2	09RYK3 deinococcus
5	36	85.7	1110	10	09SDA5 Osdas arabidopsis
6	35	83.3	306	4	09NVM2 O9nvm2 homo sapien
7	35	83.3	317	2	09XAN8 O9xan8 streptomyce
8	35	83.3	321	2	005848 O05848 mycobacteri
9	35	83.3	331	4	09UGR5 O9ugr5 homo sapien
10	35	83.3	332	4	09UGR6 O9ugr6 homo sapien
11	35	83.3	353	4	09Y515 O9y515 homo sapien
12	35	83.3	411	2	095146 P95146 mycobacteri
13	35	83.3	508	2	09KYL3 O9kyl3 streptomyce
14	35	83.3	792	4	092845 O92845 homo sapien
15	35	83.3	1114	5	09Y180 O9y180 drosophila
16	35	83.3	3198	5	026639 O26639 strongyloce
17	35	83.3	4340	2	030764 O30764 streptomyce
18	34	81.0	283	5	044989 O44989 caenorhabdi
19	34	81.0	294	5	019979 O19979 caenorhabdi

20	34	81.0	576	10	09XIC4 O9xic4 arabidopsis
21	34	81.0	1518	13	091AR8 O91ar8 salmo salar
22	34	81.0	1519	13	09W750 O9w750 salmo salar
23	34	81.0	1737	11	09J104 O9j104 ratius norv
24	34	81.0	1739	11	09J112 O9j112 mus musculu
25	34	81.0	1745	4	09N206 O9n206 homo sapien
26	34	81.0	3643	2	09N206 O9n206 streptomyce
27	33	78.6	300	5	022183 O22183 caenorhabdi
28	33	78.6	450	2	09Z8B9 O9z8b9 chlamydia p
29	33	78.6	467	2	09K296 O9k296 chlamydia p
30	33	78.6	572	5	09W327 O9w327 drosophila
31	33	78.6	589	4	013896 O13896 homo sapien
32	33	78.6	600	2	P74569 P74569 synecocyst
33	33	78.6	709	6	09N281 O9n281 canis famli
34	33	78.6	879	2	P71076 P71076 bacillus su
35	33	78.6	1007	10	09M285 O9m285 arabidopsis
36	33	78.6	1147	11	09QY21 O9qy21 mus musculu
37	33	78.6	1167	11	09JLC4 O9jlc4 mus musculu
38	33	78.6	1391	13	057539 O57539 xenopus lae
39	33	78.6	1633	13	090941 O90941 gallus gall
40	33	78.6	1832	2	09L8C8 O9l8c8 sorangium c
41	33	78.6	1832	2	09K128 O9k128 sorangium c
42	33	78.6	1838	4	015094 O15094 homo sapien
43	33	78.6	1838	4	088207 O88207 mus musculu
44	33	78.6	1840	11	060467 O60467 cricetus
45	33	78.6	1840	11	09J103 O9j103 ratius norv

ALIGNMENTS

RESULT	1				
ID	006452	PRELIMINARY:	PRT:	812 AA.	
AC	006452;				
DT	01-NOV-1996 (TREMBLrel. 01, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)				
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)				
DE	EXTRACELLULAR MATRIX PROTEIN EMFI ALPHA.				
GN	COLF1.				
OS	Ephydattia muelleri (Mueller's freshwater sponge).				
OC	Eukaryota; Metazoa; Porifera; Demospongiae; Ctenophora;				
OC	Haplosterida; Spongiillidae; Ephydattia.				
OX	NCBI_TaxID=6052;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-94047120; PubMed-8230249;				
RA	Exposito J.Y., Van der Rest M., Garrone R.;				
RT	"The complete intron/exon structure of Ephydattia muelleri fibrillar collagen gene suggests a mechanism for the evolution of an ancestral RT gene module."				
RL	J. Mol. Evol. 37:254-259(1993).				
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.				
CC	EMBL: X69818; CAA49472.1; -				
DR	INTERPRO: IPR000087; -				
DR	PFAM: PF01391; Collagen; 13.				
KW	Extracellular matrix.				
SQL	SEQUENCE 812 AA; 72280 MW; 326573F37E46D50F CRC64;				
Query Match					
Best Local Similarity 90.5%; Score 38; DB 5; Length 812;					
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
OY	1 VGVLCRPG 8				
DB	203 VGVLCRPG 210				
RESULT	2				
ID	097409	PRELIMINARY:	PRT:	189 AA.	
AC	097409;				

DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-AUG-1999 (TREMBLrel. 11, Last annotation update)
 GN SG3 OR SG3.
 GN SG3 OR SG3.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 OC Culicidae; Anopheles.
 OC NCBI_TaxID=7165;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN-GASUA; TISSUE-SALIVARY GLAND;
 RA Arca B., Lombardo F., Capurro de Lara Guimaraes M., della Torre A.,
 RT Dimopoulos G., James A.A., Coluzzi M.;
 RT "Trapping cDNAs encoding secreted proteins from the salivary glands of
 the malaria vector Anopheles gambiae.";
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN-GASUA; TISSUE-SALIVARY GLAND;
 RA Arca B., Lombardo F., de Lara Capurro M., della Torre A.,
 RT Dimopoulos G., James A.A., Coluzzi M.;
 RT "Trapping cDNAs encoding secreted proteins from the salivary glands of
 the malaria vector Anopheles gambiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1516-1521(1999).
 DR EMBL: AJ130951; CAI10260.1; -
 KW Signal.
 FT SIGNAL.
 KM
 SQ SEQUENCE 189 AA; 20039 MW; D801A07B784646FC CRC64;

Query Match
 Best Local Similarity 88.1%; Score 37; DB 5; Length 189;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VGVLRPG 8
 DB 73 VGVLRPG 80

RESULT 3
 ID 033954 PRELIMINARY; PRT; 4472 AA.
 AC 033954;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE TYLACTONE SYNTHASE STARTER MODULE AND MODULES 1 & 2.
 GN TYLG.
 OS Streptomyces fradiae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1906;
 RN 111
 RP SEQUENCE FROM N.A.
 RC Dehoff B.S., Sutton K.L., Rosteck P.R. Jr.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U78289; AAB6504.1; -
 DR INTERPRO: IPR00255; -
 DR INTERPRO: IPR00794; -
 DR INTERPRO: IPR001227; -
 DR INTERPRO: IPR001993; -
 DR INTERPRO: IPR002106; -
 DR INTERPRO: IPR002198; -
 DR PFAM: PF00106; adh_short; 1.
 DR PFAM: PF00109; ketocacyl-synt; 3.
 DR PFAM: PF00550; pp-binding; 3.
 DR PFAM: PF00698; Acyl_transf; 3.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; 2.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.

DR PROSITE: PS00339; AA_TRAN_LIGASE_II-2; UNKNOWN_3.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE: PS50075; ACP_DOMAIN; 3.
 KW Transferase
 SQ SEQUENCE 4472 AA; 463281 MW; E91D1B9E0C332CF CRC64;

Query Match
 Best Local Similarity 88.1%; Score 37; DB 2; Length 4472;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VGVLRPG 8
 DB 4155 GILGRPG 4161

RESULT 4
 ID 09RYK3 PRELIMINARY; PRT; 577 AA.
 AC 09RYK3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE LONG-CHAIN FATTY ACID--COA LIGASE.
 GN DRA0309.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Delnooccus group; Deinococcales; Deinococcus.
 OC NCBI_TaxID=1299;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN-RI;
 RX MEDLINE-20036986; PubMed-10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.S., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
 radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001863; AAF12469.1; -
 DR TIGR: DRA0309; -
 DR INTERPRO: IPR000873; -
 DR PFAM: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; UNKNOWN_1.
 KW Ligase.
 SQ SEQUENCE 577 AA; 63408 MW; C560E6D794E8826F CRC64;

Query Match
 Best Local Similarity 85.7%; Score 36; DB 2; Length 577;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VGVLRPG 7
 DB 495 VGVLRPG 501

RESULT 5
 ID 09SDA5 PRELIMINARY; PRT; 1110 AA.
 AC 09SDA5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE PURININE DISEASE RESISTANCE PROTEIN.
 GN AT2G17060.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosid II;
 OC Brassicales; Brassicaceae; Arabidopsis.

```

OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=CV. COLUMBIA;
RA  Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA  Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA  Buell C.R., Ketchum K.A., Lee J.T., Ronning C.M., Koo H., Moffat K.S.,
RA  Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA  Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA  Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA  Salzberg S.L., Fraser C.M., Venter J.C.;
RT  "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL  Nature 402:761-768(1999).
DR  EMBL: AC002354; AAF18600.1; -.
DR  INTERPRO: IPR000157; -.
DR  INTERPRO: IPR000767; -.
DR  INTERPRO: IPR001611; -.
DR  INTERPRO: IPR002182; -.
DR  Pfam: PF00560; LRR: 4;
DR  Pfam: PF00931; NB-ARC: 1.
DR  Pfam: PF01582; TIR: 1.
DR  PRINTS: PR00364; DISEASERESIST.
SQ  SEQUENCE 1110 AA; 125882 MW; 741A3E841DB9979 CRC64;

Query Match
Best Local Similarity 85.7%; Score 36; DB 10; Length 1110;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY  1 VGVLCRPG 8
    |||||
DB  147 VGVLCMPG 154

RESULT 6
ID  09NVM2 PRELIMINARY; PRT; 306 AA.
AC  09NVM2.
DT  01-OCT-2000 (TREMBlrel. 15, Created)
DT  01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE  01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE  CDNA FLJ10645 FIS, CLONE NT2RP2005767, MODERATELY SIMILAR TO G. GALLUS
DE  PRL GENE.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA  Mishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA  Tanai H., Kimata M., Watanabe M., Hirooka S., Ishii S., Kawai Y.,
RA  Saito K., Yamamoto U., Wakamatsu A., Nakamura Y., Nagahari K.,
RA  Masuko Y., Kanehori K.;
RT  "NEO human cDNA sequencing project.";
RT  Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AK001507; BAA91728.1; -.
DR  EMBL: AK001507; BAA91728.1; -.
SQ  SEQUENCE 306 AA; 33028 MW; 649E417571F59E64 CRC64;

Query Match
Best Local Similarity 83.3%; Score 35; DB 4; Length 306;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY  1 VGVLCRPG 8
    |||||
DB  178 VGVLCMPG 185

RESULT 7
ID  09XAN8 PRELIMINARY; PRT; 317 AA.
AC  09XAN8.

```

```

DT  01-NOV-1999 (TREMBlrel. 12, Created)
DT  01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT  01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE  PUTATIVE INTEGRAL MEMBRANE PROTEIN.
GN  SC4C6.02C.
OS  Streptomyces coelicolor.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX  NCBI_TaxID=1902;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=A3(2);
RA  Seeger S., Harris D.;
RT  "A set of ordered cosmids and a detailed genetic and physical map for
RT  the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RT  Submitted (Jun-1999) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=A3(2);
RA  James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL  Submitted (Jun-1999) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=A3(2);
RX  MEDLINE=97000351; PubMed=8643436;
RA  Redenbach M., Kieseer H.M., Denapalte D., Eichner A., Cullum J.,
RA  Kinshl H., Hopwood D.A.;
RT  "A set of ordered cosmids and a detailed genetic and physical map for
RT  the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL  Mol. Microbiol. 21:77-96(1996).
DR  EMBL: AL079355; CAB45566.1; -.
SQ  SEQUENCE 317 AA; 30829 MW; 3372F5FB132062B9 CRC64;

Query Match
Best Local Similarity 83.3%; Score 35; DB 2; Length 317;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY  1 VGVLCRPG 8
    |||||
DB  142 IGVLSRPG 149

RESULT 8
ID  005848 PRELIMINARY; PRT; 321 AA.
AC  005848;
DT  01-JUL-1997 (TREMBlrel. 04, Created)
DT  01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT  01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE  HYPOHETICAL 34.5 KDA PROTEIN.
GN  RV3218 OR MTCY07D11.08C.
OS  Mycobacterium tuberculosis.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1773;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=H37RV;
RX  MEDLINE=98295987; PubMed=9634230;
RA  Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA  Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA  Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA  Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA  Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA  Rutter S., Seeger K., Skellon S., Squares S., Squires R., Sulston J.E.,
RA  Taylor K., Whitehead S., Barrell B.G.;
RT  "Deciphering the biology of Mycobacterium tuberculosis from the
RT  complete genome sequence.";
RL  Nature 393:537-544(1998).
DR  EMBL: Z95120; CAB08319.1; -.
DR  TUBERCULIST; RV3218; -.

```

DR INTERPRO: IPR001206; -
 DR PRAM: PR00781; DAGKc; 1.
 DR PRODOM: PD005043; -; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 321 AA; 34474 MW; 164014C4944CAC9 CRC64;

Query Match
 Best Local Similarity 83.3%; Score 35; DB 2; Length 321;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLGRPG 8
 Db 73 GMLGRPG 79

RESULT 9
 ID 09UGR5 PRELIMINARY; PRT; 331 AA.
 AC 09UGR5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE DJ889J22B.1 (NOVEL PROTEIN (ISOFORM 1)) (FRAGMENT).
 GN DJ889J22B.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Steward C.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL049851; CAB63076.1; -
 FT NON_TER 1
 SQ SEQUENCE 331 AA; 36925 MW; DFF0A6323E39F65E CRC64;

Query Match
 Best Local Similarity 83.3%; Score 35; DB 4; Length 331;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLGRPG 8
 Db 255 GLLGRPG 261

RESULT 10
 ID 09UGR6 PRELIMINARY; PRT; 332 AA.
 AC 09UGR6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE DJ889J22B.1 (NOVEL PROTEIN (ISOFORM 2)) (FRAGMENT).
 GN DJ889J22B.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Steward C.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL049851; CAB63075.1; -
 FT NON_TER 1
 SQ SEQUENCE 332 AA; 37053 MW; 68234A527D009007 CRC64;

Query Match
 Best Local Similarity 83.3%; Score 35; DB 4; Length 332;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVLGRPG 8
 Db 256 GLLGRPG 262

RESULT 11
 ID 09Y515 PRELIMINARY; PRT; 353 AA.
 AC 09Y515;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE DJ682J15.1 (NOVEL COLLAGEN TRIPLE HELIX REPEAT CONTAINING PROTEIN)
 GN DJ682J15.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Smith S.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL034452; CAB45235.1; -
 DR INTERPRO: IPR000087; -
 DR PFM: PFM01391; Collagen; 4.
 FT NON_TER 1
 SQ SEQUENCE 353 AA; 34746 MW; 1E331F6D0B93DEBB CRC64;

Query Match
 Best Local Similarity 83.3%; Score 35; DB 4; Length 353;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVGRPG 8
 Db 253 VGVGRPG 260

RESULT 12
 ID P95146 PRELIMINARY; PRT; 411 AA.
 AC P95146;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DE HYPOHETICAL 43.6 KDA PROTEIN.
 GN RV1869C OR MTC1359.04.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN NCBI_TaxID=1773;
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Davies R., Devlin K., Brown D., Chillingworth T., Connor R.,
 RA Hornsby T., Jagels K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Oliver S., Osborne J., Quail M.A., McLean J., Moule S., Murphy L.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL: Z83859; CAB06118.1; -
 DR TUBERCULIST: RV1869C; -
 DR INTERPRO: IPR000205; -
 DR INTERPRO: IPR001100; -
 DR PFM: PFM0070; Pyl_redox; 1.

KW Hypothetical protein.
SQ SEQUENCE 411 AA; 43629 MW; 26643CA800C8D879 CRC64;

Query Match

Best Local Similarity 83.3%; Score 35; DB 2; Length 411;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVGLRPG 8
Db 311 GVGLRPG 317

RESULT 13

ID 09KYL3 PRELIMINARY; PRT; 508 AA.
AC 09KYL3;
DT 01-OCT-2000 (TREMBLREL. 15, Created)
DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)
DE NITRATE REDUCTASE (FRAGMENT).
GN NARB.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinsahl H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL356592; CAB92217.1; -;
FT NON-TER 1
SQ SEQUENCE 508 AA; 55055 MW; 8E138F2612E4EDDB CRC64;

Query Match 83.3%; Score 35; DB 2; Length 508;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVGLRPG 8
Db 35 GVGLRPG 41

RESULT 14

ID 092845 PRELIMINARY; PRT; 792 AA.
AC 092845;
DT 01-FEB-1997 (TREMBLREL. 02, Created)
DT 01-FEB-1997 (TREMBLREL. 02, Last sequence update)
DE SMAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;

RX MEDLINE=97059159; PubMed=8900189;

RA Shimizu K., Kawabe H., Minami S., Honda T., Takatsuki K., Shirataki H.,
RA Takai Y.;

RT "SMAP, an Smg GDS-associating protein having arm repeats and
RT phosphorylated by Src tyrosine kinase.";

RL J. Biol. Chem. 271:27013-27017(1996).

DR EMBL; U59919; AAC50788.1; -;

SQ INTERPRO; IPR000225; -;
SEQUENCE 792 AA; 91189 MW; 07367E27FC092E5D CRC64;

Query Match 83.3%; Score 35; DB 4; Length 792;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GVGLRPG 7
Db 768 GVGLRPG 774

RESULT 15

ID 09V180 PRELIMINARY; PRT; 1114 AA.
AC 09V180;
DT 01-MAY-2000 (TREMBLREL. 13, Created)
DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
DE CG14608 PROTEIN.
GN CG14608.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyloroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchem K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson R.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
DR EMBL: AE003672; AAF54045.1; -;
DR FLYBASE: FBgn0037487; CG14608.
DR INTERPRO: IPR001395; -;
DR PROSITE: PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
SQ SEQUENCE 1114 AA; 122174 MW; BA30C7512EFD4B3 CRC64;

Query Match 83.3%; Score 35; DB 5; Length 1114;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GVLGRPG 8
11:1111
Db 67 GVGGRPG 73

Search completed: June 13, 2001, 14:29:46
Job time: 547 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw.model

Run on: June 13, 2001, 14:27:07 ; Search time 78.71 Seconds
(without alignments)
1.953 Million cell updates/sec

Title: PCT-US01-05825A-29

Perfect score: 42

Sequence: 1 VGVLRPG 8

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCRNUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	88.1	4472	2 US-08-804-227C-2	Sequence 2, Appli
2	35	83.3	4545	2 US-08-804-227C-14	Sequence 14, Appli
3	35	83.3	4550	2 US-08-804-227C-8	Sequence 8, Appli
4	35	83.3	4550	2 US-08-804-198-2	Sequence 2, Appli
5	34	81.0	785	1 US-08-526-964-4	Sequence 4, Appli
6	34	81.0	785	1 US-08-946-617-4	Sequence 4, Appli
7	34	81.0	785	1 US-09-031-897-4	Sequence 4, Appli
8	33	78.6	8	3 US-08-689-421-6	Sequence 6, Appli
9	33	78.6	8	4 US-09-389-528-6	Sequence 6, Appli
10	33	78.6	1832	4 US-09-335-409-4	Sequence 4, Appli
11	33	78.6	2500	2 US-08-801-263A-2	Sequence 2, Appli
12	33	78.6	2500	3 US-09-102-248-2	Sequence 2, Appli
13	33	78.6	2512	2 US-08-801-263A-9	Sequence 9, Appli
14	33	78.6	2512	2 US-09-102-248-9	Sequence 9, Appli
15	33	78.6	2517	2 US-08-801-263A-5	Sequence 5, Appli
16	33	78.6	2517	3 US-09-102-248-5	Sequence 5, Appli
17	32	76.2	584	2 US-08-415-593-41	Sequence 41, Appli
18	31	73.8	21	3 US-08-851-843A-175	Sequence 175, App
19	31	73.8	21	4 US-08-974-549A-294	Sequence 294, App
20	31	73.8	111	4 US-08-466-886-19	Sequence 19, Appli
21	31	73.8	111	4 US-08-469-617-19	Sequence 19, Appli
22	31	73.8	309	1 US-08-463-082B-8	Sequence 8, Appli
23	31	73.8	309	2 US-08-460-907B-8	Sequence 8, Appli
24	31	73.8	549	1 US-08-494-168-8	Sequence 8, Appli
25	31	73.8	684	1 US-08-555-669-12	Sequence 12, Appli
26	31	73.8	684	4 US-09-073-663-12	Sequence 12, Appli
27	31	73.8	1479	2 US-08-951-912-4	Sequence 4, Appli

28	31	73.8	1480	1 US-07-637-621-2	Sequence 2, Appli
29	31	73.8	1480	1 US-08-136-742A-2	Sequence 2, Appli
30	31	73.8	1480	1 US-08-135-809A-2	Sequence 2, Appli
31	31	73.8	1480	1 US-08-466-886-17	Sequence 17, Appli
32	31	73.8	1480	2 US-08-951-912-2	Sequence 2, Appli
33	31	73.8	1480	2 US-08-951-912-6	Sequence 6, Appli
34	31	73.8	1480	2 US-08-469-461-2	Sequence 2, Appli
35	31	73.8	1480	2 US-08-463-461-4	Sequence 4, Appli
36	31	73.8	1480	2 US-08-691-605-2	Sequence 2, Appli
37	31	73.8	1480	2 US-08-455-552A-14	Sequence 14, Appli
38	31	73.8	1480	3 US-07-890-609-2	Sequence 2, Appli
39	31	73.8	1480	3 US-07-890-609-4	Sequence 4, Appli
40	31	73.8	1480	3 US-09-248-026-2	Sequence 2, Appli
41	31	73.8	1480	4 US-08-463-617-17	Sequence 17, Appli
42	31	73.8	1480	5 PCT-US93-11667-2	Sequence 2, Appli
43	31	73.8	1480	6 PCT-US93-11667-2	Sequence 2, Appli
44	31	73.8	1694	1 US-08-494-168-2	Sequence 2, Appli
45	30	71.4	9	4 US-08-660-092-126	Sequence 126, App

ALIGNMENTS

RESULT 1

US-08-804-227C-2

Sequence 2, Application US/08804227C

Patent No. 5676991

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Kustoss, Stuart A.

APPLICANT: Rostock, Paul R., Jr.

APPLICANT: Sutton, Kimberly L.

TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESSES:

ADDRESSEE: THOMAS G. PLANT 1501

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCII(DOS) Text only

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,227C

FILING DATE: February 21, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.

REGISTRATION NUMBER: 35,784

REFERENCE/DOCKET NUMBER: X-8231

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-2459

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 4472 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-804-227C-2

Query Match

Best Local Similarity 88.1%; Score 37; DB 2; Length 4472;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VGVLRPG 8

DB 4155 GILGRG 4161

```

RESULT 2
US-08-804-227C-14
; Sequence 14, Application US/08804227C
; Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kustoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804, 227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4545 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-14

Query Match
Best Local Similarity 83.3%; Score 35; DB 2; Length 4545;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLGRPG 8
Db 4218 GILGRPG 4224

RESULT 3
US-08-804-227C-8
; Sequence 8, Application US/08804227C
; Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kustoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804, 227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-8

Query Match
Best Local Similarity 83.3%; Score 35; DB 2; Length 4550;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLGRPG 8
Db 4223 GILGRPG 4229

RESULT 4
US-08-804-198-2
; Sequence 2, Application US/08804198
; Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kustoss, Stuart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rostock, Paul R., Jr.
TITLE OF INVENTION: PLANTENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804, 198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cantrell, Paul R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-2
```


Query Match 83.3%; Score 35; DB 2; Length 4550;
Best Local Similarity 85.7%; Pred. No. 6.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLGRPG 8
|:|||||
Db 4223 GILGRPG 4229

RESULT 5

US-08-526-964-4
; Sequence 4, Application US/08526964
; Patent No. 5698421
; GENERAL INFORMATION:
; APPLICANT: Lamdowitz, Alan M
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; APPLICANT: Yang, Jian
; TITLE OF INVENTION: Nucleotide Integrase Preparation
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter & Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,964
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Golrick, Mary E
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8200
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 785 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
US-08-526-964-4

Query Match 81.0%; Score 34; DB 1; Length 785;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLGRPG 8
|:|||||
Db 534 GILGRPG 540

RESULT 6

US-08-946-617-4
; Sequence 4, Application US/08946617
; Patent No. 5869634
; GENERAL INFORMATION:
; APPLICANT: Lamdowitz, Alan M
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; APPLICANT: Yang, Jian
; TITLE OF INVENTION: Nucleotide Integrase Preparation

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:

ADDRESSEE: Calfee, Halter & Griswold
STREET: 800 Superior Avenue
CITY: Cleveland
STATE: Ohio
COUNTRY: USA
ZIP: 44114

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/946,617

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Golrick, Mary E

REGISTRATION NUMBER: 34,829

REFERENCE/DOCKET NUMBER: 22727/00127

TELECOMMUNICATION INFORMATION:

TELEPHONE: (216) 622-8200

TELEFAX: (216) 241-0816

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 785 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHEICAL: NO

US-08-946-617-4

Query Match 81.0%; Score 34; DB 2; Length 785;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLGRPG 8
|:|||||
Db 534 GILGRPG 540

RESULT 7

US-09-031-897-4
; Sequence 4, Application US/09031897
; Patent No. 6027895
; GENERAL INFORMATION:
; APPLICANT: Lamdowitz, Alan
; APPLICANT: Mohr, Georg
; APPLICANT: Zimmerly, Steven
; APPLICANT: Guo, Huatao
; TITLE OF INVENTION: Methods Cleaving DNA with Nucleotide
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter & Griswold
; STREET: 800 Superior Avenue, Suite 1400
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,897
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

NAME: Docherty, Pamela A.
REGISTRATION NUMBER: 40,591
REFERENCE/DOCKET NUMBER: 24671/00105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216)622-8416
TELEFAX: (216)241 0816
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 785 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-09-031-897-4

Query Match 81.0%; Score 34; DB 3; Length 785;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLGPG 8
|:|:|:|
Db 534 GILGKPG 540

RESULT 8
US-08-689-421-6
Sequence 6, Application US/08689421
Patent No. 6008029
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberley M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halikier, Torben P
TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6008029 No. 6008029disk of No. 6008029th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/689,421
FILING DATE: 9-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4554, 204-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-689-421-6

Query Match 78.6%; Score 33; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 1.4e+05;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GVLGPG 8
|:|:|:|
Db 1 GVLGNPG 7

RESULT 9
US-09-389-528-6
Sequence 6, Application US/09389528
Patent No. 6207430
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberley M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halikier, Torben P
TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6207430 No. 6207430disk of No. 6207430th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/389,528
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/689,421
FILING DATE: 9-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4554, 204-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-389-528-6

Query Match 78.6%; Score 33; DB 4; Length 8;
Best Local Similarity 85.7%; Pred. No. 1.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVLGPG 8
|:|:|:|
Db 1 GVLGNPG 7

RESULT 10
US-09-335-409-4
Sequence 4, Application US/09335409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross

APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHLONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1832
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-4

Query Match 78.6%; Score 33; DB 4; Length 1832;
Best Local Similarity 75.0%; Pred. No. 6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VGVLRPG 8
Db 1578 VGLGSPG 1585

RESULT 11
US-08-801-263A-2
Sequence 2, Application US/08801263A
Patent No. 5811407
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5811407th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,263A
FILING DATE: 19-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2500 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-263A-2

Query Match 78.6%; Score 33; DB 2; Length 2500;
Best Local Similarity 62.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 VGVLRPG 8
:||:| |

Db 722 IGVIGTGP 729

RESULT 12
US-09-102-248-2
Sequence 2, Application US/09102248
Patent No. 6008035
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6008035th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,248
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2500 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-102-248-2

Query Match 78.6%; Score 33; DB 3; Length 2500;
Best Local Similarity 62.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VGVLRPG 8
Db 722 IGVIGTGP 729

RESULT 13
US-08-801-263A-9
Sequence 9, Application US/08801263A
Patent No. 5811407
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte

STATE: NO. 5811407ch Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,263A
FILING DATE: 19-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2512 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-263A-9

Query Match 78.6%; Score 33; DB 2; Length 2512;
Best Local Similarity 62.5%; Pred. No. 8.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVIGRPG 8
Db 722 IGVIGTPG 729

RESULT 14
US-09-102-248-9
Sequence 9, Application US/09102248
Patent No. 6008035
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 6008035th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,248
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/801,263
FILING DATE: 19-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2512 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-102-248-9

Query Match 78.6%; Score 33; DB 3; Length 2512;
Best Local Similarity 62.5%; Pred. No. 8.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVIGRPG 8
Db 722 IGVIGTPG 729

RESULT 15
US-08-801-263A-5
Sequence 5, Application US/08801263A
Patent No. 5811407
GENERAL INFORMATION:
APPLICANT: Johnston, Robert E.
APPLICANT: Davis, Nancy L.
APPLICANT: Simpson, Dennis A.
TITLE OF INVENTION: System for the In Vivo Delivery and
TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATE: No. 5811407ch Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,263A
FILING DATE: 19-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-147
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2517 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-263A-5

Query Match 78.6%; Score 33; DB 2; Length 2517;
Best Local Similarity 62.5%; Pred. No. 8.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VGVIGRPG 8
Db 722 IGVIGAPG 729

Search completed: June 13, 2001, 14:27:08

Wed Jun 13 15:00:43 2001

pct-us01-05825a-29.ra1

Page 7

Job time: 629 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:25:48 ; Search time 150.28 Seconds
(without alignments)
3.043 Million cell updates/sec

Title: PCT-US01-05825A-30

Perfect score: 43

Sequence: 1 VDCGRIG 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_0401:*

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	97.7	22	16	R85657 Mastitis vaccine p
2	42	97.7	25	16	R85659 Mastitis vaccine m
3	38	88.4	8	21	V79129 Peptide antagonist
4	38	88.4	12	18	W07623 N-terminal peptide
5	38	88.4	12	20	W94929 US864028 sequence
6	38	88.4	18	18	W27314 Human prostate ass
7	38	88.4	18	19	W45521 Glyceraldehyde-3-p
8	38	88.4	18	20	W02457 Fragment of glycer
9	38	88.4	22	21	V79132 Human foetal brain
10	38	88.4	124	20	V29539 Human lung tumour
11	38	88.4	124	21	B44446 Human lung tumour

12	38	88.4	127	20	V29540 Human lung tumour
13	38	88.4	127	21	B44447 Human lung tumour
14	38	88.4	268	20	V14921 Amino acid sequenc
15	38	88.4	335	20	Y07036 Breast cancer asso
16	38	88.4	335	20	Y05368 Human HCMV Inducib
17	38	88.4	340	20	Y14930 Amino acid sequenc
18	37	86.0	31	19	W53450 Bifidobacterium ge
19	37	86.0	39	14	R38788 Streptococcal surf
20	37	86.0	47	21	G15911 Arabidopsis thalia
21	37	86.0	70	14	R40212 N-terminal of a pl
22	37	86.0	93	21	G34299 Arabidopsis thalia
23	37	86.0	107	21	G19017 Zea mays protein f
24	37	86.0	130	21	G44842 Zea mays protein f
25	37	86.0	133	21	G33938 Arabidopsis thalia
26	37	86.0	135	21	G22593 Zea mays protein f
27	37	86.0	136	21	G54490 Zea mays protein f
28	37	86.0	139	18	W27835 Arabidopsis thalia
29	37	86.0	149	21	G35523 Arabidopsis thalia
30	37	86.0	156	20	Y06889 S. aureus glyceral
31	37	86.0	160	18	W14722 C. hitusulus GPD.
32	37	86.0	171	21	G22592 Zea mays protein f
33	37	86.0	174	21	G35522 Arabidopsis thalia
34	37	86.0	175	21	G28442 Arabidopsis thalia
35	37	86.0	238	21	B28132 Glyceraldehyde-3-p
36	37	86.0	333	19	W55089 Streptococcus pneu
37	37	86.0	333	21	B19020 Amino acid sequenc
38	37	86.0	333	21	B19021 Amino acid sequenc
39	37	86.0	334	11	R03211 Amino acid sequenc
40	37	86.0	334	21	Y55844 P. ciferrii glycer
41	37	86.0	335	11	R05284 Amino acid sequenc
42	37	86.0	335	15	Y28472 Glyceraldehyde-3-p
43	37	86.0	336	15	R56486 Plasmin receptor.
44	37	86.0	336	20	Y35244 C. pneumoniae prot
45	37	86.0	336	22	Y85681 Streptococcal plas

ALIGNMENTS

RESULT 1	
ID R85657	standard; peptide: 22 AA.
XX	
AC R85657:	
XX	
DT 19-JAN-1996	(first entry)
XX	
DE Mastitis vaccine peptide.	
XX	
KW Mastitis vaccine peptide; microbial adherence.	
XX	
OS Synthetic.	
XX	
FH	
FT Modified-site	Location/Qualifiers
FT	1 /note= "N-terminal hydrogen or C1-C16
FT	carboxylic acid"
FT	
XX	
Modified-site	22 /note= "C-terminal hydroxy group"
XX	
PN W09512410-A1.	
XX	
PD 11-MAY-1995.	
XX	
PF 04-NOV-1994:	94WO-US12752.
XX	
PR 05-NOV-1993:	93US-0147765.
XX	
PA (ELIL) LILLY & CO ELI.	
XX	
PI Scheffinger CC, Smiley DL;	
XX	
DR WPI, 1995-185598/24.	

Query Match	97.7%	Score 42;	DB 16;	Length 25;
Best Local Similarity	87.5%;	Pred. No. 0.15;		
Matches	7;	Conservative	1;	Mismatches
			0;	Indels
				Gaps
0;				
QY	1	VDSFGRRIG	8	
Db	6	ldgfgrrig	13	
RESULT	3			
ID	Y79129			
XX	Y79129	standard; Peptide; 8 AA.		
AC	Y79129;			

Accession	Source	Protein	Protein description
U00000	NCBI	3	RESULT
U00000	NCBI	Y79129	Y79129 standard; Peptide; 8 AA.
U00000	NCBI	Y79129;	
U00000	NCBI	05-JUN-2000	(first entry)
U00000	NCBI	Peptide antagonist of zonulin.	
U00000	NCBI	Zonulin; antagonist; zonula occludens toxin receptor;	
U00000	NCBI	blood-brain barrier; antiinflammatory; cerebroprotective;	
U00000	NCBI	neuroprotective; dermatological; antiulcer; antiviral;	
U00000	NCBI	antibacterial; cytostatic; anti-HIV; vulvener; antiallergic;	
U00000	NCBI	hypotensive; immunosuppressive; antiparasitic; vasotropic;	
U00000	NCBI	brain injury; therapy.	
U00000	NCBI	Synthetic.	
U00000	NCBI	WO200007609-A1.	
U00000	NCBI	17-FEB-2000.	
U00000	NCBI	28-JUL-1999; 99WO-US16683.	
U00000	NCBI	03-AUG-1998; 98US-0127815.	
U00000	NCBI	(UYMA-) UNIV MARYLAND BALTIMORE.	
U00000	NCBI	Fasano A;	
U00000	NCBI	WPI; 2000-205565/18.	
U00000	NCBI	New peptide antagonist of zonulin useful as antiinflammatory agent for treating cerebral ischemia.	

PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 XX shigellosis, viral gastroenteritis, meningitis, encephalomyelitis
 PS Claim 1; Page 52, 69pp; English.
 XX
 CC This present sequence is that of a peptide antagonist of zornin,
 CC one of 25 such peptides (see Y79105-20) of the invention that
 CC bind to a zonula occludens toxin (ZOT) receptor, yet do not
 CC physiologically modulate the opening of mammalian tight junctions
 CC (TJ). The peptide antagonists are based on a common motif of ZOT
 CC and human zonulins, which is believed to be critical for receptor
 CC binding. They can be prepared by chemical synthesis or by use of
 CC recombinant DNA techniques. The present peptide antagonist is
 CC used in a claimed method for treatment of a condition associated
 CC with breakdown of the blood-brain barrier, where it binds to the
 CC ZOT receptor in the brain but does not modulate the opening of TJ
 CC in the brain. It can be used for treating a condition associated
 CC with breakdown of the blood-brain barrier such as osmotic injury,
 CC e.g. cerebral ischemia, stroke or cerebral oedema, hypertension,
 CC convulsive seizure, chemical toxins, uraemia, meningitis,
 CC encephalitis, encephalomyelitis, e.g. infective, or bacterial or
 CC allergic, tumors, traumatic brain injuries, radiation brain injury,
 CC immaturity and kernicterus, demyelinating diseases, e.g. multiple
 CC sclerosis or Guillain-Barre syndrome.
 XX
 Sequence 8 AA:

Query Match 88.4%: Score 38; DB 21; Length 8;
 Best Local Similarity 87.5%: Pred. NO. 3.2e+05;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGFRIG 8
 1:|||||
 Db 5 vngfgriq 12

RESULT 4
 W07623
 ID W07623 standard; peptide; 12 AA.
 AC W07623;
 DT 08-SEP-1997 (first entry)
 DE N-terminal peptide of GAPDH.
 XX
 KW Ribozyme; tumour necrosis factor-alpha; rheumatic arthritis; AIDS;
 KW autoimmune disease; TNF-alpha; viral disease.
 XX
 OS Synthetic.
 XX
 PN W09639499-A1.
 XX
 PD 12-DEC-1996.
 XX
 PE 05-JUN-1996; 96WO-AU00339.
 XX
 PR 05-JUN-1995; 95US-0464073.
 XX
 PA (GENE-) GENE SHEARS PTY LTD.
 XX
 PI Sloud M;
 XX
 DR WPI: 1997-043121/04.
 XX
 PT Tumour necrosis factor-alpha ribozymes - useful to treat disorders
 associated with TNF-alpha overexpression, e.g. rheumatic arthritis,
 PT AIDS or autoimmune disease
 XX
 PS Example 2; Page 61; 158pp; English.
 XX
 CC This sequence represents the N-terminal peptide of GAPDH. GAPDH
 CC binds to a sequence found in anti-tumour necrosis factor-alpha
 CC (TNF-alpha) ribozymes. The binding of GAPDH acts to stabilise the
 CC ribozyme and has the ability to stimulate cleavage activity of
 CC hammerhead ribozymes. This increase is most likely due to the
 CC destabilisation of the RNA duplex.
 XX
 SO Sequence 12 AA;

Query Match 88.4%: Score 38; DB 18; Length 12;
 Best Local Similarity 87.5%: Pred. NO. 0.39;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGFRIG 8
 1:|||||
 Db 5 vngfgriq 12

RESULT 5
 W94929
 ID W94929 standard; peptide; 12 AA.
 AC W94929;
 DT 06-MAY-1999 (first entry)
 DE
 XX

DE US5864028 sequence ID #33.
 XX
 KW TNF-alpha; ribozyme; growth hormone; blood factor; enzyme; vaccine;
 KW virus; tumour necrosis factor; rheumatoid arthritis; septic shock;
 KW acquired immune deficiency syndrome; graft vs. host disease; cachexia;
 KW immune dysfunction; Alzheimer's disease; psoriasis; leukemia; cancer;
 KW mRNA stabilisation.
 XX
 OS Unidentified.
 XX
 PN US5864028-A.
 XX
 PD 26-JAN-1999.
 XX
 PE 22-JUN-1995; 95US-0428252.
 XX
 PR 22-JUN-1995; 95US-0428252.
 PR 03-NOV-1992; 92US-0971058.
 PR 03-NOV-1993; 93WO-AU00567.
 PR 05-JUN-1995; 95US-0464073.
 XX
 PA (GENE-) GENE SHEARS PTY LTD.
 XX
 PI Sloud M;
 XX
 DR WPI: 1999-131361/11.
 XX
 PT Stabilising RNA by attachment to specific second RNA sequence - for
 improving activity of ribozymes or antisense molecules or for
 PT improving polypeptide production from mRNA
 XX
 PS Disclosure: Column 51-52; 79pp; English.
 XX
 CC The invention relates to a composition that comprises a first RNA
 CC covalently linked to second RNA which includes at least a sequence shown
 CC in X05901. Attachment of the second RNA is used to stabilise the first
 CC RNA, which may be a ribozyme or antisense molecule, or mRNA encoding a
 CC polypeptide, e.g. a growth hormone, blood factor, enzyme, or antigen for
 CC vaccine. The ribozymes are particularly directed against viruses
 CC (pathogenic in animals or plants), tumour necrosis factor-alpha (TNF-
 CC alpha) for treating e.g. rheumatoid arthritis, acquired immune deficiency
 CC syndrome, septic shock, graft vs. host diseases, and cachexia or designed
 CC to treat a wide variety of other diseases such as immune dysfunction,
 CC Alzheimer's disease, psoriasis, and leukemia and other cancers.
 CC Stabilising mRNA with the second RNA improves the production of proteins,
 CC particularly in animal cells, substantially reducing costs, and increases
 CC the effect of ribozymes or antisense molecules.
 XX
 SO Sequence 12 AA;

Query Match 88.4%: Score 38; DB 20; Length 12;
 Best Local Similarity 87.5%: Pred. NO. 0.39;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDFGFRIG 8
 1:|||||
 Db 5 vngfgriq 12

RESULT 6
 W27314
 ID W27314 standard; peptide; 18 AA.
 AC W27314;
 DT 27-APR-1998 (first entry)
 DE Human prostate associated peptide.
 XX
 DE Human prostate associated peptide.
 XX
 KW Human; prostate cancer; immunotherapy; therapy; immunodiagnosis;
 KW diagnosis; vaccine; antibody; human.
 XX

PT diagnosis and monitoring of prostate cancer
 XX Claim 12; Page 88; 106pp; English.
 PS
 CC The present sequence represents the N-terminal of a human
 CC glyceraldehyde-3-phosphate dehydrogenase. The specification describes
 CC polypeptides and DNA which are obtained from prostate tumour
 CC cells. The polypeptide comprises an immunogenic portion of a
 CC prostate protein. The polypeptides and DNA obtained from prostate
 CC tumour cells, as well as antibodies raised against the protein, can
 CC be used in the treatment, diagnosis and monitoring of prostate
 CC cancer.
 CC
 SQ Sequence 18 AA:
 OY 1 VDGFGRIG 8
 DB 7 vngfgrrig 14
 RESULT 9
 ID Y79132 standard; Peptide; 22 AA.
 XX Y79132;
 AC Y79132;
 XX 05-JUN-2000 (first entry)
 DE Human foetal brain zonulin N-terminal sequence.
 XX
 DE Zonulin; antagonist; zonula occludens toxin receptor;
 KW human; blood-brain barrier; antiinflammatory;
 KW gastrointestinal inflammation; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note="unidentified residue"
 XX
 FT WO200007609-A1.
 XX
 PN 17-FEB-2000.
 PD
 XX 28-JUL-1999; 99WO-US16683.
 PE
 XX 03-AUG-1998; 98US-0127815.
 PR
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX
 PI Fasano A;
 XX
 DR WPI: 2000-205565/18.
 DR
 XX New peptide antagonist of zonulin useful as antiinflammatory agent for
 PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
 PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis -
 PT
 XX Example 4; Fig 7; 69pp; English.
 PS
 XX The present sequence is that of the N-terminal region of foetal
 CC human brain zonulin. The N-terminal sequences of human adult and
 CC foetal zonulins (see Y79130-36) were compared with *Vibrio cholerae*
 CC zonula occludens toxin (ZOT) to identify a common motif thought
 CC to be involved in receptor binding. Peptide antagonists (see
 CC Y79105-29) based on this motif are useful as antiinflammatory
 CC agents for treatment of gastrointestinal inflammation, and for
 CC treatment of conditions associated with breakdown of the blood-brain

CC barrier.
 XX Sequence 22 AA:
 SQ
 OY Query Match 88.4%; Score 38; DB 21; Length 22;
 DB Best Local Similarity 87.5%; Pred. No. 0.73;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VDGFGRIG 8
 DB 8 vngfgrrig 15
 RESULT 10
 ID Y29539 standard; Protein; 124 AA.
 XX Y29539;
 AC Y29539;
 XX 13-OCT-1999 (first entry)
 DT
 XX Human Lung tumour protein L86S-49 predicted amino acid sequence.
 DE Human Lung tumour protein L86S-49 predicted amino acid sequence.
 XX
 KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
 KW immunotherapy; detection; inhibition.
 KW
 OS Homo sapiens.
 XX
 OS WO938973-A2.
 PN
 XX 05-AUG-1999.
 PD
 XX 26-JAN-1999; 99WO-US01642.
 PE
 XX 22-DEC-1998; 98US-0219245.
 PR 28-JAN-1998; 98US-0015022.
 PR 28-JAN-1998; 98US-0015022.
 PR 18-MAR-1998; 98US-0040828.
 PR 18-MAR-1998; 98US-0040831.
 PR 23-JUL-1998; 98US-0122191.
 PR 23-JUL-1998; 98US-0122192.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
 PI
 XX WPI: 1999-479187/40.
 DR N-PSDB; 207223.
 DR
 XX Lung tumour specific polynucleotides for inhibiting the development
 PT of lung cancer
 PT
 XX Example 3; Page 114; 171pp; English.
 PS
 XX The present invention describes lung tumour specific polynucleotides
 CC and tumour antigens. Z07144 to Z07246 and Z08301 to Z08325 represent
 CC specifically claimed polynucleotides, and Z29486 to Z29571 represent
 CC amino acid sequences from the present invention. The lung tumour
 CC specific polynucleotides and polypeptides can be used in pharmaceutical
 CC compositions and vaccines to inhibit the development of lung cancer.
 CC They can also be used to detect lung cancer in a patient. Probes and
 CC antibodies derived from the lung tumour sequences are useful in
 CC detection of lung cancer.
 CC
 SQ Sequence 124 AA:
 OY Query Match 88.4%; Score 38; DB 20; Length 124;
 DB Best Local Similarity 87.5%; Pred. No. 4.2;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VDGFGRIG 8

```

Db      8 vngfgrrg 15

RESULT 11
ID      B44446 standard; Protein; 124 AA.
XX
AC      B44446;
XX
DT      05-FEB-2001 (first entry)
XX
DE      Human lung tumour-specific antigen encoded by cDNA #57.
XX
KM      Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
OS      Homo sapiens.
XX
PN      WO200060077-A2.
XX
PD      12-OCT-2000.
XX
PE      30-MAR-2000; 2000WO-US08560.
XX
PR      02-APR-1999; 99US-0285323.
XX
PR      09-AUG-1999; 99US-0370838.
XX
PR      30-DEC-1999; 99US-0476235.
XX
PR      03-MAR-2000; 2000US-0518809.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Reed SG, Lodes MJ, Mohamath R, Secrlist H;
XX
DR      WPI; 2000-638466/61.
XX
DR      N-PSDB; C79104.
XX
PT      Novel lung tumor polypeptides and polynucleotides, useful for
XX
PT      detecting, monitoring or treating cancer, especially lung cancer -
XX
PS      Claim 1; Page 132; 243pp; English.
XX
XX
CC      The present sequence is given in a specification relating to compounds
CC      for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC      least an immunogenic part of a lung tumour protein are disclosed.
CC      The polypeptides are useful for inhibiting the development of cancer,
CC      especially lung cancer. Samples of T cells expressing the polypeptides
CC      may be used to inhibit the development of cancer. The polypeptides are
CC      also useful for detecting and monitoring the progression of cancer,
CC      especially lung cancer.
XX
SQ      Sequence 124 AA;

Query Match      88.4%; Score 38; DB 21; Length 124;
Best Local Similarity 87.5%; Pred. No. 4.2;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 VDGEGRIG 8
      1:|||||
Db      8 vngfgrrg 15

RESULT 12
ID      Y29540 standard; Protein; 127 AA.
XX
AC      Y29540;
XX
DT      13-OCT-1999 (first entry)
XX
DE      Human lung tumour protein L86S-51 predicted amino acid sequence.
XX
KM      Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
XX

```

```

KM      immunotherapy; detection; inhibition.
XX
OS      Homo sapiens.
XX
PN      WO938973-A2.
XX
PD      05-AUG-1999.
XX
PE      26-JAN-1999; 99WO-US01642.
XX
PR      22-DEC-1998; 98US-0219245.
XX
PR      28-JAN-1998; 98US-0015022.
XX
PR      28-JAN-1998; 98US-0015029.
XX
PR      18-MAR-1998; 98US-0040828.
XX
PR      18-MAR-1998; 98US-0040831.
XX
PR      23-JUL-1998; 98US-0122191.
XX
PR      23-JUL-1998; 98US-0122192.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
XX
DR      WPI; 1999-479187/40.
XX
DR      N-PSDB; 207224.
XX
PT      Lung tumour specific polynucleotides for inhibiting the development
XX
PT      of lung cancer
XX
PS      Example 3; Page 114-115; 171pp; English.
XX
XX
CC      The present invention describes lung tumour specific polynucleotides
CC      and tumour antigens. 207144 to 207246 and 208301 to 208325 represent
CC      specifically claimed polynucleotides, and Y29486 to Y29571 represent
CC      amino acid sequences from the present invention. The lung tumour
CC      specific polynucleotides and polypeptides can be used in pharmaceutical
CC      compositions and vaccines to inhibit the development of lung cancer.
CC      They can also be used to detect lung cancer in a patient. Probes and
CC      antibodies derived from the lung tumour sequences are useful in
CC      detection of lung cancer.
XX
SQ      Sequence 127 AA;

Query Match      88.4%; Score 38; DB 20; Length 127;
Best Local Similarity 87.5%; Pred. No. 4.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 VDGEGRIG 8
      1:|||||
Db      26 vngfgrrg 33

RESULT 13
ID      B44447 standard; Protein; 127 AA.
XX
AC      B44447;
XX
DT      05-FEB-2001 (first entry)
XX
DE      Human lung tumour-specific antigen encoded by cDNA #58.
XX
KM      Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
OS      Homo sapiens.
XX
PN      WO200060077-A2.
XX
PD      12-OCT-2000.
XX
PE      30-MAR-2000; 2000WO-US08560.
XX
PR      02-APR-1999; 99US-0285323.
XX

```

PR 09-AUG-1999; 99US-0370838.
 PR 30-DEC-1999; 99US-0476235.
 PR 03-MAR-2000; 2000US-0518809.
 XX
 PA (CORI-) CORIXA CORP.
 PI Reed SG, Lodes MJ, Mohamath R, Secrist H;
 DR WPI: 2000-638466/61.
 DR N-PSDB; C79105.
 XX
 PT Novel lung tumor polypeptides and polynucleotides, useful for
 PT detecting, monitoring or treating cancer, especially lung cancer -
 PS
 PS Claim 1; Page 132; 243pp; English.
 XX
 CC The present sequence is given in a specification relating to compounds
 CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
 CC least an immunogenic part of a lung tumour protein are disclosed.
 CC The polypeptides are useful for inhibiting the development of cancer,
 CC especially lung cancer. Samples of T cells expressing the polypeptides
 CC may be used to inhibit the development of cancer. The polypeptides are
 CC also useful for detecting and monitoring the progression of cancer,
 CC especially lung cancer.
 CC
 SQ Sequence 127 AA;

Query Match 88.4%; Score 38; DB 21; Length 127;
 Best Local Similarity 87.5%; Pred. No. 4.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRTG 8
 I:|||||
 Db 26 vngfgrrlg 33

RESULT 14

Y14921
 ID Y14921 standard; protein: 268 AA.

AC Y14921;

DT 25-OCT-1999 (first entry)

DE Amino acid sequence of M. vaccae antigen GV-44.

XX
 KW Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
 KW dendritic cell maturation; infectious disease; immune disorder; cancer;
 KW respiratory system; mycobacterial infection; allergy; tuberculosis;
 KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
 KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 KW squamous cell carcinoma; melanoma.

OS Mycobacterium vaccae.

PN WO9932634-A2.

PD 01-JUL-1999.

PF 23-DEC-1998; 98WO-NZ00189.

PR 04-DEC-1998; 98US-0205426.

PR 23-DEC-1997; 97US-0996624.

PR 23-DEC-1997; 97US-0997080.

PR 11-JUN-1998; 98US-0095855.

PR 17-SEP-1998; 98US-0156181.

PA (GENE-) GENESTIS RES & DEV CORP LTD.
 PI Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
 XX

DR WPI: 1999-430163/36.
 DR N-PSDB; Z11385.

XX Enhancing Immune response to an antigen

PS Claim 1; Page 231-232; 243pp; English.

XX
 CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
 CC T cells and natural killer cells, to stimulate the production of
 CC cytokines, to enhance the expression of co-stimulatory molecules on
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation
 CC and function. The proteins can be expressed by standard recombinant
 CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma.
 CC
 SQ Sequence 268 AA;

Query Match 88.4%; Score 38; DB 20; Length 268;
 Best Local Similarity 87.5%; Pred. No. 9.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRTG 8
 I:|||||
 Db 7 vngfgrrlg 14

RESULT 15

Y07036
 ID Y07036 standard; Protein: 335 AA.

AC Y07036;

DT 02-JUL-1999 (first entry)

DE Breast cancer associated antigen precursor sequence.

XX
 KW Cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; lung cancer;
 KW prostate cancer.

OS Homo sapiens.

PN WO9904265-A2.

PD 28-JAN-1999.

PF 15-JUL-1998; 98WO-US14679.

PR 22-JUN-1998; 98US-0102322.

PR 17-JUL-1997; 97US-0896164.

PR 10-OCT-1997; 97US-0061599.

PR 10-OCT-1997; 97US-0061765.

PR 10-OCT-1997; 97US-0948705.

PR 11-OCT-1997; 97GB-0021697.

PA (LUDWIG-) LUDWIG INST CANCER RES.

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
 PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
 PI Tureci O;
 DR WPI: 1999-132448/11.
 XX

PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers

XX Disclosure; Page 421-422; 787pp; English.

XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.

XX Sequence 335 AA:

Query Match 88.4%; Score 38; DB 20; Length 335;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGGRIG 8
|:|||||
Db 8 vngfgrig 15

Search completed: June 13, 2001, 14:25:49
Job time: 668 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2001, 14:23:12 ; Search time 87.97 Seconds
(without alignments)
6.250 Million cell updates/sec

Title: PCT-US01-05825a-30

Perfect score: 43

Sequence: 1 VDFGFRIG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.67.*

1: pirl.*
2: pirl2.*
3: pirl3.*
4: pirl4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	332	1	DEPG3 glyceraldhyde-3-p
2	42	97.7	333	1	DELOG3 glyceraldhyde-3-p
3	40	93.0	67	2	A19822 glyceraldhyde-3-p
4	39	90.7	25	2	S62673 glyceraldhyde-3-p
5	38	88.4	21	2	B12055 glyceraldhyde-3-p
6	38	88.4	34	2	A12055 glyceraldhyde-3-p
7	38	88.4	333	1	DEHYG glyceraldhyde-3-p
8	38	88.4	333	1	DEMSG glyceraldhyde-3-p
9	38	88.4	333	1	DERTG glyceraldhyde-3-p
10	38	88.4	333	1	DECHG3 glyceraldhyde-3-p
11	38	88.4	333	1	DEKZGR glyceraldhyde-3-p
12	38	88.4	333	2	UC4309 glyceraldhyde-3-p
13	38	88.4	333	2	JC5370 glyceraldhyde-3-p
14	38	88.4	333	1	JN0678 glyceraldhyde-3-p
15	38	88.4	335	1	DEHUG3 glyceraldhyde-3-p
16	37	86.0	22	2	B60475 glyceraldhyde-3-p
17	37	86.0	24	2	PN0163 glyceraldhyde-3-p
18	37	86.0	27	2	PA0083 glyceraldhyde-3-p
19	37	86.0	31	2	S13205 glyceraldhyde-3-p
20	37	86.0	31	2	S13206 glyceraldhyde-3-p
21	37	86.0	34	2	A60475 glyceraldhyde-3-p
22	37	86.0	165	2	S14477 glyceraldhyde-3-p
23	37	86.0	329	1	DEVKGL glyceraldhyde-3-p
24	37	86.0	330	2	E75408 glyceraldhyde-3-p
25	37	86.0	331	1	DEECG3 glyceraldhyde-3-p
26	37	86.0	331	1	DEUTGC glyceraldhyde-3-p
27	37	86.0	331	1	DETWG3 glyceraldhyde-3-p
28	37	86.0	331	2	S57279 glyceraldhyde-3-p
29	37	86.0	331	2	S57280 glyceraldhyde-3-p

30	37	86.0	331	2	S57281 glyceraldhyde-3-p
31	37	86.0	331	2	B48445 glyceraldhyde-3-p
32	37	86.0	332	1	F82131 glyceraldhyde-3-p
33	37	86.0	332	1	DEJNG1 glyceraldhyde-3-p
34	37	86.0	332	1	DEBYG1 glyceraldhyde-3-p
35	37	86.0	332	1	DEBYG2 glyceraldhyde-3-p
36	37	86.0	332	1	DEBYG3 glyceraldhyde-3-p
37	37	86.0	332	2	A22366 glyceraldhyde-3-p
38	37	86.0	332	2	B22366 glyceraldhyde-3-p
39	37	86.0	332	2	JC4373 glyceraldhyde-3-p
40	37	86.0	332	2	S24630 glyceraldhyde-3-p
41	37	86.0	332	2	I40069 glyceraldhyde-3-p
42	37	86.0	332	2	G71879 glyceraldhyde-3-p
43	37	86.0	332	2	A46635 glyceraldhyde-3-p
44	37	86.0	332	2	C81285 glyceraldhyde-3-p
45	37	86.0	333	1	DEHGR glyceraldhyde-3-p

ALIGNMENTS

RESULT 1
DEPG3
glyceraldhyde-3-phosphate dehydrogenase (EC 1.2.1.12) - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 08-Oct-1981 #sequence_revision 08-Oct-1981 #text_change 31-Dec-1996
C:Accession: A93156; A94413; A00367
R:Harris, J.I.; Perham, R.N.
Nature 219, 1025-1028, 1968
A:Title: Glyceraldhyde 3-phosphate dehydrogenase from pig muscle.
A:Reference number: A93156; MUID:68399311
A:Accession: A93156
A:Molecule type: Protein
A:Residues: 1-44, 'E', '46-332 <HAR>
R:Harris, J.I.; Davidson, B.E.; Saigo, M.; Noller, H.F.; Perham, R.N.
In Enzymes and Isoenzymes: Structure, Properties and Function, Shugar, D., ed., pp.1-
A:Reference number: A94413
A:Accession: A94413
A:Molecule type: Protein
A:Residues: 45 <HAR>
C:Comment: Cys-149 covalently binds glyceraldhyde-3-phosphate.
C:Superfamily: glyceraldhyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F:2-32/Region: beta-alpha-beta NAD nucleotide-binding fold
F:149,176/Active site: Cys, His #status predicted

Query Match 100.0%; Score 43; DB 1; Length 332;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDFGFRIG 8
DB 5 VDFGFRIG 12

RESULT 2
DELOG3
glyceraldhyde-3-phosphate dehydrogenase (EC 1.2.1.12) - American lobster
C:Species: Homarus americanus (American lobster)
C:Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 31-Dec-1996
C:Accession: A00369
R:Davidson, B.E.; Saigo, M.; Noller, H.F.; Harris, J.I.
Nature 216, 1181-1185, 1967
A:Title: Amino-acid sequence of glyceraldhyde 3-phosphate dehydrogenase from lobster
A:Reference number: A93153; MUID:66098010
A:Accession: A00369
A:Molecule type: Protein
A:Residues: 1-333 <DAV>
A:Note: The source species is not provided by the authors
R:Moras, D.; Olsen, K.W.; Sabesan, M.N.; Buehner, M.; Ford, G.C.; Rossmann, M.G.
J Biol Chem 250, 9137-9162, 1975
A:Title: Studies of asymmetry in the three-dimensional structure of lobster D-glyceral

A:Reference number: A92188; MUID:76069246
 A:Contents: annotation; X-ray crystallography, 2.9 angstroms
 A:Note: the source species is not provided by the authors
 R:Allison, W.S.; Kaplan, N.O.
 J. Biol. Chem. 239, 2140-2152, 1964
 A>Title: The comparative enzymology of triosephosphate dehydrogenase.
 A:Reference number: A44655
 A:Contents: annotation; source
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: acetylated amino end; gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
 F:2-3/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:1/Modified site: acetylated amino end (Ser) #status experimental
 F:148,175/Active site: Cys, His #status experimental

Query Match 97.7%; Score 42; DB 1; Length 333;
 Best Local Similarity 87.5%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRIG 8
 :|||||||
 Db 5 IDGFGRIG 12

RESULT 3
 A19822
 glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - Bacillus coagulans (tentative S
 C:Species: Bacillus coagulans
 C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
 C:Accession: A19822; A05100
 R:Crabb, J.W.; Murdock, A.L.; Suzuki, T.; Hamilton, J.W.; McLinden, H.; Amelunxen, R.E.
 J. Bacteriol. 143, 503-512, 1981
 A>Title: Sequence homology in the amino-terminal and active-site regions of thermostable
 A:Reference number: A91788; MUID:81117013
 A:Accession: A19822
 A:Molecule type: protein
 A:Residues: 1-67 <CRA>
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match 93.0%; Score 40; DB 2; Length 67;
 Best Local Similarity 75.0%; Pred. No. 0.61;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRIG 8
 :|||||||
 Db 6 IDGFGRIG 13

RESULT 4
 S62673
 glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12), hibernating tissue - desert jerb
 C:Species: Jaculus orientalis (desert jerboa)
 C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 19-Feb-1999
 C:Accession: S62673
 R:Soukri, A.; Hafid, N.; Valverde, F.; Elkebbaj, M.S.; Serrano, A.
 Biochim. Biophys. Acta 1292, 177-187, 1996
 A>Title: Evidence for a posttranslational covalent modification of liver glyceraldehyde-
 A:Reference number: S62673; MUID:96139342
 A:Accession: S62673
 A:Molecule type: protein
 A:Residues: 1-25 <SOU>
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; hibernation; homotetramer; NAD; oxidoreductase

Query Match 90.7%; Score 39; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DGFGRIG 8
 :|||||||

Db 6 DGFGRIG 12

RESULT 5
 B12055
 glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - pig (fragments)
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 12-Apr-1995
 C:Accession: B12055
 R:Kulbe, K.D.; Jackson, K.W.; Tang, J.
 Biochem. Biophys. Res. Commun. 67, 35-42, 1975
 A>Title: Structural evidence for a liver-specific glyceraldehyde-3-phosphate dehydrog
 A:Reference number: A94641; MUID:76087882
 A:Accession: B12055
 A:Molecule type: protein
 A:Residues: 1-21 <KUL>
 A:Experimental source: liver
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match 88.4%; Score 38; DB 2; Length 21;
 Best Local Similarity 87.5%; Pred. No. 0.46;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDGFGRIG 8
 :|||||||
 Db 5 VDNFGRIG 12

RESULT 6
 A12055
 glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - bovine (fragments)
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 05-Jun-1987 #sequence_revision 30-Sep-1993 #text_change 11-May-2000
 C:Accession: A12055; A36630; S03922; A60790
 R:Kulbe, K.D.; Jackson, K.W.; Tang, J.
 Biochem. Biophys. Res. Commun. 67, 35-42, 1975
 A>Title: Structural evidence for a liver-specific glyceraldehyde-3-phosphate dehydrog
 A:Reference number: A94641; MUID:76087882
 A:Accession: A12055
 A:Molecule type: protein
 A:Residues: 1-30 <KUL>
 A:Experimental source: liver
 R:Hsu, S.C.; Molday, R.S.
 J. Biol. Chem. 265, 13308-13313, 1990
 A>Title: Glyceraldehyde-3-phosphate dehydrogenase is a major protein associated with
 A:Reference number: A36630; MUID:90330679
 A:Accession: A36630
 A:Molecule type: protein
 A:Residues: 1-22 <HSU>
 A:Experimental source: rod outer segment
 R:Kocha, T.; Fukuda, T.; Isobe, T.; Okuyama, T.
 Biochim. Biophys. Acta 991, 56-61, 1989
 A>Title: Specific purification of glyceraldehyde-3-phosphate dehydrogenase by hydroph
 A:Reference number: S03922; MUID:89229261
 A:Accession: S03922
 A:Molecule type: protein
 A:Residues: 1-15 <KOC>
 R:Krawczyk, E.; Broda, K.; Sidorowicz, A.; Golebiewska, J.; Stenienlowski, H.; Banas,
 Comp. Biochem. Physiol. B 85, 811-818, 1986
 A>Title: Comparative study of the structure of glyceraldehyde-3-phosphate dehydrogena
 A:Reference number: A60790; MUID:87132300
 A:Accession: A60790
 A:Molecule type: protein
 A:Residues: 1-7,31-34 <KRA>

A:Experimental source: heart muscle
 A:Note: the authors suggest that this protein differs from the form in liver
 A:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase


```

Query Match      88.4%: Score 38: DB 2: Length 34:
Best Local Similarity 87.5%: Pred. No. 0.75:
Matches 7: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

OY 1 VDGFRIG 8
DB 5 VNGFRIG 12

RESULT 7
DEHYG
glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - Chinese hamster
C:Species: Crictetus griseus (Chinese hamster)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 11-Jun-1999
C:Accession: S10221
R:Vincent, S.; Fort, P.
Nucleic Acids Res. 18, 3054, 1990
A:Title: Nucleotide sequence of hamster glyceraldehyde-3-phosphate dehydrogenase mRNA.
A:Reference number: S10221; MUID:90272420
A:Accession: S10221
A:Molecule type: mRNA
A:Residues: 1-333 <VIN>
A:Cross-references: EMBL:X51213; NID:949434; PIDN:CAA36368.1; PID:949435
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
F:150/177/Active site: Cys, His #status predicted

Query Match      88.4%: Score 38: DB 1: Length 333:
Best Local Similarity 87.5%: Pred. No. 7.3:
Matches 7: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

OY 1 VDGFRIG 8
DB 6 VNGFRIG 13

RESULT 8
DEMSG
glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999
C:Accession: J00553; S14160
R:Sabat, D.E.; Broome, H.E.; Prystowsky, M.B.
Gene 91, 185-191, 1990
A:Title: Glyceraldehyde-3-phosphate dehydrogenase mRNA is a major interleukin 2-induced
A:Reference number: J00553; MUID:91007274
A:Accession: J00553
A:Molecule type: mRNA
A:Residues: 1-333 <SAB>
A:Cross-references: GB:M2559; NID:9193423; PIDN:AAA37659.1; PID:9309243
R:Filipek, A.; Gerke, V.; Weber, K.; Kuznicki, J.
Eur. J. Biochem. 195, 795-800, 1991
A:Title: Characterization of the cell-cycle-regulated protein calyculin from Ehrlich asc
A:Reference number: S14090; MUID:91153321
A:Accession: S14160
A:Molecule type: protein
A:Residues: 2-20, 'FSCD', '25-26', 'D', '28-64', 'I', '66-70', 'E', '72-77', 'F', '79-80', 'VK', '83-86', 'D', '88-
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F:2-33/Region: beta-alpha-beta NAD nucleotide-binding fold
F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
F:150/177/Active site: Cys, His #status predicted

Query Match      88.4%: Score 38: DB 1: Length 333:
Best Local Similarity 87.5%: Pred. No. 7.3:
Matches 7: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

OY 1 VDGFRIG 8
DB 6 VNGFRIG 13

RESULT 9
DEPTG
glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - rat
N:Alternate names: triosephosphate dehydrogenase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jun-1999
C:Accession: A23013; A22939; A22887; A23280; A17155; J00401; A60208; B17155
R:Fort, P.; Marty, L.; Plechaczky, M.; El Sabouty, S.; Danl, C.; Jeanneur, P.; Blanc
Nucleic Acids Res. 13, 1431-1442, 1985
A:Title: Various rat adult tissues express only one major mRNA species from the glyce
A:Reference number: A23013; MUID:85215556
A:Accession: A23013
A:Molecule type: mRNA
A:Residues: 1-333 <FOR>
A:Cross-references: GB:X02231; NID:956187; PIDN:CAA26150.1; PID:956188
R:Tso, J.Y.; Sun, X.H.; Kao, T.; Reece, K.S.; Wu, R.
Nucleic Acids Res. 13, 2485-2502, 1985
A:Title: Isolation and characterization of rat and human glyceraldehyde-3-phosphate d
A:Reference number: A93562; MUID:85215629
A:Accession: A22939
A:Molecule type: mRNA
A:Residues: 1-80, 'AN', '83-304', 'F', '306-333 <TNSO>
A:Cross-references: GB:M17701; NID:9204248; PIDN:AAA41193.1; PID:9204249
R:Plechaczky, M.; Blanchard, J.M.; Marty, L.; Danl, C.; Panabieres, F.; El Sabouty, S
Nucleic Acids Res. 12, 6951-6963, 1984
A:Title: Post-transcriptional regulation of glyceraldehyde-3-phosphate-dehydrogenase
A:Reference number: A22887; MUID:85014145
A:Accession: A22887
A:Molecule type: mRNA
A:Residues: 261-323 <PIE>
A:Cross-references: GB:X00972
R:Maehara, Y.; Fujiyoshi, T.; Takahashi, K.; Yamamoto, M.; Endo, H.
Biochem. Biophys. Res. Commun. 131, 800-805, 1985
A:Title: 1.5 KB mRNA abundantly expressed in rat tumors encodes a 37 kilodalton prote
A:Reference number: A23280; MUID:86025533
A:Accession: A23280
A:Molecule type: mRNA
A:Residues: 267-304, 'F', '306-333 <MAE>
A:Cross-references: GB:M11561; NID:9205963; PIDN:AAA1795.1; PID:9205964
R:Vospelinkova, N.D.; Sifonova, M.I.; Shuvalova, E.R.; Baratova, L.A.; Kniazev, S.P.
Biochem. J. 199, 757-765, 1981
A:Title: Identification of an arginine residue important for catalytic activity in th
A:Reference number: A90313; MUID:82182080
A:Accession: A17155
A:Molecule type: protein
A:Residues: 117-119, 'N', '121-122', 'T', '124-126', 'Z', '128', 'LF', '131-133', 'BRZH', '138', 'SK', '294-
R:Barbakov, B.A.; Zheltova, A.O.; Belyanova, L.P.; Baratova, L.A.; Vospelinkova, N.D.
Bioorg. Khim. 3, 826-830, 1977
A:Title: Peptide sequence containing the active site cysteine of D-glyceraldehyde-3-p
A:Reference number: J00401
A:Accession: J00401
A:Molecule type: protein
A:Residues: 144-160 <BAT>
A:Experimental source: skeletal muscle
R:Leung, T.K.C.; Hall, C.; Konifiles, C.; Lim, L.
J. Neurochem. 49, 232-238, 1987
A:Title: Tirifluperazine activates and releases latent ATP-generating enzymes associa
A:Reference number: A60208; MUID:87224934
A:Accession: A60208
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 235-304, 'F', '306-333 <LEU>
A:Cross-references: GB:M29341; NID:9203141; PIDN:AAA40814.1; PID:9203142
A:Experimental source: brain
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
F:150/Active site: Cys #status predicted
F:177/Active site: His #status predicted

```

Query Match 88.4%; Score 38; DB 1; Length 333;
 Best Local Similarity 87.5%; Pred. No. 7.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDNFGRIG 8
 I:|||||
 DB 6 VDNFGRIG 13

RESULT 10
 DECHG3
 glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - chicken

N:Alternate names: GAP dehydrogenase
 C:Species: Gallus gallus (chicken)
 C:Date: 03-Aug-1984 #sequence_revision 30-Sep-1989 #text_change 16-Jun-2000
 C:Accession: A00368; A32737; A22035; I50231; I50640
 R:Duigalczky, A.; Haron, J.A.; Stone, E.M.; Dennison, O.E.; Rothblum, K.N.; Schwartz, R.J.
 Biochemistry 22, 1605-1613, 1983
 A:Title: Cloning and sequencing of a deoxyribonucleic acid copy of glyceraldehyde-3-phos-
 A:Reference number: A00368; MUID:83204759
 A:Accession: A00368
 A:Molecule type: mRNA
 A:Residues: 1-333 <DNUG>
 A:Cross-references: GB:J00849; NID:963402; PIDN:CAA23698.1; PID:g1628381
 R:Panabieres, F.; Plechaczky, M.; Ratner, B.; Danl, C.; Fort, P.; Rlaad, S.; Marty, L.;
 Biochem. Biophys. Res. Commun. 118, 767-773, 1984
 A:Title: Complete nucleotide sequence of the messenger RNA coding for chicken muscle gly-
 A:Reference number: A32737; MUID:8415834
 A:Accession: A32737
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-333 <PNAN>
 A:Cross-references: GB:K01458; NID:9211800; PIDN:AAA48778.1; PID:g211801
 R:Stone, E.M.; Rothblum, K.N.; Alewy, M.C.; Kuo, T.M.; Schwartz, R.J.
 Proc. Natl. Acad. Sci. U.S.A. 82, 1628-1632, 1985
 A:Title: Complete sequence of the chicken glyceraldehyde-3-phosphate dehydrogenase gene
 A:Reference number: A22035; MUID:85166184
 A:Accession: A22035
 A:Molecule type: DNA
 A:Residues: 1-293, 'H', 295-333 <STO>
 A:Cross-references: GB:M11213; NID:g211796; PIDN:AAA48774.1; PID:g211797
 A:Note: the authors translated the codon CAT for residue 294 as Asp
 J. Arnold, H.H.; Domdey, H.; Wlebauner, K.; Datta, K.; Siddiqui, M.A.O.
 J. Biol. Chem. 257, 9872-9877, 1982
 A:Title: Cloning, partial sequencing, and expression of glyceraldehyde-3-phosphate dehy-
 A:Reference number: I50231; MUID:82265644
 A:Accession: I50231
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'G', 198-276, 'E', 278-333 <ARN>
 A:Cross-references: GB:J00848; NID:g211798; PIDN:AAA48777.1; PID:g211799
 A:Accession: I50640
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 'G', 198-276, 'E', 278-328, 'T', 330-333 <AR2>
 A:Cross-references: EMBL:V00406; NID:963400; PIDN:CAA23697.1; PID:963401
 C:Genetics: 8/2; 41/3; 77/3; 107/3; 146/2; 173/3; 224/2; 250/3; 278/3; 311/2
 A:Introns: 9/2; 41/3; 77/3; 107/3; 146/2; 173/3; 224/2; 250/3; 278/3; 311/2
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconogenesis; glycolysis; homotetramer; NAD; oxidoreductase
 F:2-333/Product: glyceraldehyde-3-phosphate dehydrogenase #status predicted <MAT>
 F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:150,177/Active site: Cys, His #status predicted

Query Match 88.4%; Score 38; DB 1; Length 333;
 Best Local Similarity 87.5%; Pred. No. 7.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDNFGRIG 8
 I:|||||
 DB 6 VDNFGRIG 13

RESULT 11

DEKZ8R

glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - yeast (Zygosaccharomyces rou-
 C:Species: Zygosaccharomyces rouxii, Candida mogii
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
 C:Accession: S00152
 R:Imura, T.; Utatsu, I.; Toh-e, A.
 Agric. Biol. Chem. 51, 1641-1647, 1987
 A:Title: Glyceraldehyde-3-phosphate dehydrogenase genes of Zygosaccharomyces rouxii:
 A:Reference number: S00152
 A:Accession: S00152
 A:Molecule type: DNA
 A:Residues: 1-333 <IMC>
 A:Cross-references: EMBL:D00134; NID:9218528; PIDN:BA00081.1; PID:9218529
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: gluconogenesis; glycolysis; homotetramer; NAD; oxidoreductase
 F:3-33/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:150,177/Active site: Cys, His #status predicted

Query Match 88.4%; Score 38; DB 1; Length 333;
 Best Local Similarity 87.5%; Pred. No. 7.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDNFGRIG 8
 I:|||||
 DB 6 VDNFGRIG 13

RESULT 12

JC4309

glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - rabbit
 N:Alternate names: triose phosphate dehydrogenase
 N:Contains: uracil DNA glycosylase (EC 3.2.2.-), nuclear
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 11-Jun-1999
 C:Accession: JC4309; I46482
 R:Applquist, S.E.; Keyna, U.; Calvin, M.R.; Beck-Engesser, G.B.; Raman, C.; Jaeck, H.
 Gene 163, 325-326, 1995
 A:Title: Sequence of the rabbit glyceraldehyde-3-phosphate dehydrogenase-encoding cDN
 A:Reference number: JC4309; MUID:96011658
 A:Accession: JC4309
 A:Molecule type: mRNA
 A:Residues: 1-333 <APP>
 A:Cross-references: GB:I23961; NID:9406106; PIDN:AAA5218.1; PID:9406107
 A:Experimental source: spleen
 R:Putney, S.D.; Herlihy, W.C.; Schimmel, P.
 Nature 302, 718-721, 1983
 A:Title: A new tropoin T and cDNA clones for 13 different muscle proteins, found by
 A:Reference number: I46471; MUID:83167564
 A:Accession: I46482
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 33-79 <PUT>
 A:Cross-references: EMBL:V00884; NID:91563; PIDN:CAA24253.1; PID:9292756
 C:Genetics:
 A:Gene: gapdh
 C:Complex: homotetramer; crystallizes as two symmetric dimers in which the partners h
 C:Function: monomer
 C:Function: <CVT>
 A:Description: oxidoreductase; reversibly catalyzes the oxidative phosphorylation of
 A:Pathway: gluconogenesis; glycolysis
 A:Note: tetrameric form; cytosol
 C:Function: <NUC>
 A:Description: DNA repair; uracil DNA glycosylase for base-excision in DNA repair
 A:Note: monomeric form; nuclear
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: cytosol; DNA repair; gluconogenesis; glycolysis; glycosylase; homotetram

Query Match 88.4%; Score 38; DB 2; Length 333;
 Best Local Similarity 87.5%; Pred. No. 7.3;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VDFGFRIG 8
:|||||
Db 6 VNGFGRIG 13

RESULT 13

JC5370
glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12), eutheric tissue - desert jerboa
N:Alternate names: GAPDH; triosephosphate dehydrogenase
C:Species: Jaculus orientalis (desert jerboa)
C:Date: 28-May-1997 #sequence_revision 12-Sep-1997 #text_change 20-Jun-2000
R:Accession: JC5370, PC4314, S62674
R:Soukri, A.; Valverde, F.; Hatid, N.; Elkebbaj, M.S.; Serrano, A.
Gene 181, 139-145, 1996
A:Title: Occurrence of a differential expression of the glyceraldehyde-3-phosphate dehydrogenase in the liver of the desert jerboa
A:Reference number: JC5370; MUID:97128782
A:Accession: JC5370
A:Molecule type: mRNA
A:Residues: 1-333 <SCOU1>
A:Cross-references: EMBL:X87226; NID:91103589; PIDN:CAA60678.1; PID:91103590
A:Accession: PC4314
A:Molecule type: protein
A:Residues: 1-56 <SCOU2>
A:Experimental source: skeletal muscle
R:Soukri, A.; Hatid, N.; Valverde, F.; Elkebbaj, M.S.; Serrano, A.
Biochim. Biophys. Acta 1292, 177-187, 1996
A:Title: Evidence for a posttranslational covalent modification of liver glyceraldehyde-3-phosphate dehydrogenase in the liver of the desert jerboa
A:Reference number: S62673; MUID:96139342
A:Accession: S62674
A:Molecule type: protein
A:Residues: 2-26 <SCOU>
C:Genetics:
A:Gene: gapc
C:Complex: homotetramer
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; hibernation; homotetramer; oxidoreductase
F:2-333/Product: glyceraldehyde-3-phosphate dehydrogenase #status experimental <MAY>

Query Match 88.4%; Score 38; DB 2; Length 333;
Best Local Similarity 87.5%; Pred. No. 7.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VDFGFRIG 8
:|||||
Db 6 VNGFGRIG 13

RESULT 14

JN0678
glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - quail
C:Species: Coturnix coturnix (quail)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Jun-1999
R:Accession: JN0678; S35726
R:Meisler, R.; Siemester, G.; Hartl, M.; Bister, K.
Gene 128, 269-272, 1993
A:Title: Sequence and expression of a glyceraldehyde-3-phosphate dehydrogenase-encoding gene in the quail
A:Reference number: JN0678; MUID:93292997
A:Accession: JN0678
A:Molecule type: mRNA
A:Residues: 1-333 <MEI>
A:Cross-references: GB:219086; NID:962615; PIDN:CAA79512.1; PID:962616
C:Genetics:
A:Gene: GAPDH
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match 88.4%; Score 38; DB 2; Length 333;
Best Local Similarity 87.5%; Pred. No. 7.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VDFGFRIG 8
:|||||
Db 6 VNGFGRIG 13

RESULT 15

DEH03
glyceralddehyde-3-phosphate dehydrogenase (EC 1.2.1.12) [validated] - human
N:Alternate names: triose phosphate dehydrogenase
N:Contains: uracil DNA glycosylase (EC 3.2.2.-), nuclear
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 08-Dec-1994 #text_change 15-Sep-2000
R:Accession: A31988; A00365; A21939; 153309; B22939; A45924; 155258; A41297; S26758;
R:Ericolani, L.; Florence, B.; Denaro, M.; Alexander, M.
J. Biol. Chem. 263, 15335-15341, 1988
A:Title: Isolation and complete sequence of a functional human glyceraldehyde-3-phosphate dehydrogenase cDNA
A:Reference number: A31988; MUID:85008430
A:Accession: A31988
A:Molecule type: DNA
A:Residues: 1-335 <ERC>
A:Cross-references: GB:J04038; NID:9182980; PIDN:AA53191.1; PID:9182981
R:Arcari, P.; Martineili, R.; Salvatore, F.
Nucleic Acids Res. 12, 9179-9189, 1984
A:Title: The complete sequence of a full length cDNA for human liver glyceraldehyde-3-phosphate dehydrogenase
A:Reference number: A00365; MUID:85087928
A:Accession: A00365
A:Molecule type: mRNA
A:Residues: 1-224, 'D', 226-335 <ARC>
A:Cross-references: GB:X01677; GB:R03121; GB:M17851; GB:X01110; NID:931644; PIDN:CAA2
R:Hanaauer, A.; Mandel, J.L.
EMBO J. 3, 2627-2633, 1984
A:Title: The glyceraldehyde 3 phosphate dehydrogenase gene family: structure of a human cDNA
A:Reference number: A21939; MUID:85076585
A:Accession: A21939
A:Molecule type: mRNA
A:Residues: 1-335 <HAN>
A:Cross-references: GB:X01677; GB:R03121; GB:M17851; GB:X01110; NID:931644
R:Diail, C.; Plechaczky, M.; Audigier, Y.; El Sabouty, S.; Cathala, G.; Marty, L.; For
Eur. J. Biochem. 145, 299-304, 1984
A:Title: Characterization of the transcription products of glyceraldehyde 3-phosphate dehydrogenase genes
A:Reference number: 153309; MUID:85051356
A:Accession: 153309
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-310, 'V', 312-335 <DAN>
A:Cross-references: GB:M28283; NID:9182978; PIDN:AA52519.1; PID:9182979
R:Tso, J.Y.; Sun, X.H.; Kao, T.; Reece, K.S.; Wu, R.
Nucleic Acids Res. 13, 2485-2502, 1985
A:Title: Isolation and characterization of rat and human glyceraldehyde-3-phosphate dehydrogenase cDNAs
A:Reference number: A93562; MUID:85215629
A:Accession: B22939
A:Molecule type: mRNA
A:Residues: 2-335 <TSO>
A:Cross-references: GB:M17851; NID:9182860; PIDN:AA56283.1; PID:9182861
R:Tokuura, K.; Nakamura, Y.; Sakata, K.; Fujimori, K.; Okubo, M.; Sawada, K.; Sakai
Cancer Res. 47, 5616-5619, 1987
A:Title: Enhanced expression of a glyceraldehyde-3-phosphate dehydrogenase gene in human cancer cells
A:Reference number: A45924; MUID:88026722
A:Accession: A45924
A:Molecule type: mRNA
A:Residues: 1-335 <TOK>
A:Cross-references: GB:M33197; NID:9182976; PIDN:AA52518.1; PID:9182977
R:Allen, R.W.; Trach, K.A.; Hoch, J.A.
J. Biol. Chem. 262, 649-653, 1987
A:Title: Identification of the 37-kDa protein displaying a variable interaction with glyceraldehyde-3-phosphate dehydrogenase
A:Reference number: 155258; MUID:87109159
A:Accession: 155258
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-335 <ALL>
A:Cross-references: GB:J02642; NID:9182862; PIDN:AA52496.1; PID:9182863
R:Meyer-Siegler, K.; Mauro, D.J.; Seal, G.; Wurzer, J.; deitel, J.K.; Sirover, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 8460-8464, 1991
 A:Title: A human nuclear uracil DNA glycosylase is the 37-kDa subunit of glyceraledehyde-
 A:Reference number: A41297; MUID:92020872
 A:Accession: A41297
 A:Molecule type: mRNA
 A:Residues: 1-335 <MEY>
 A:Cross-references: GB:X53778
 R:Strover, M. A.
 submitted to the EMBL Data Library, July 1990
 A:Reference number: S26758
 A:Accession: S26758
 A:Molecule type: mRNA
 A:Residues: 1-193, 'NGCYMAA', 201, 203, 'SRTSLPL', 213, 'L', 215-329, 333-335 <STR>
 A:Cross-references: EMBL:X53778; NID:935052; PIDN:CA37794.1; PID:935053
 R:Nowak, K.; Kuczek, M.; Ostropolska, L.; Malarska, A.; Wolny, M.; Branowski, T.
 Hoppe-Seyler's Z. Physiol. Chem. 356, 1181-1183, 1975
 A:Title: The covalent structure of glyceraledehyde-phosphate dehydrogenase from human mus
 A:Reference number: A12103; MUID:76067491
 A:Accession: A12103
 A:Molecule type: protein
 A:Residues: 2-8, 'B', 10-23, 'B', 25-27, 46-47, 'Z', 49, 'B', 51-61, 67-69, 'B', 71, 'K', 'B', 82, 'AZ',
 'V', 171, 'ZZ', 174-176, 'SS', 179-180, 'AB', 186-188, 'B', 190-194, 'B', 199-202, 'A', 204-205, 'L', 20
 'Z', 318-321, 323-334, 'SKGK' <NOM>
 A:Note: some of this partial sequence was assigned tentatively based on composition
 R:Nowak, K.; Wolny, M.; Banas, T.
 FEBS Lett. 134, 143-146, 1981
 A:Title: The complete amino acid sequence of human muscle glyceraledehyde 3-phosphate de
 A:Reference number: A00366; MUID:82073291
 A:Accession: A00366
 A:Molecule type: Protein
 A:Residues: 2-8, 'D', 10-40, 'H', 42-63, 'D', 65-69, 'D', 71, 'KA', 74-82, 'EN', 85-91, 'TA', 94-112,
 226-242, 'L', 244-263, 'E', 265-278, 'DE', 281-283, 'D', 285-287, 'GSN', 291-293, 'I', 295-301, 'E',
 R:Sloud, M.; Jespersen, L.
 J. Mol. Biol. 257, 775-789, 1996
 A:Title: Enhancement of hammerhead ribozyme catalysis by glyceraledehyde-3-phosphate dehy
 A:Reference number: S66563; MUID:96194445
 A:Accession: S66563
 A:Molecule type: protein
 A:Residues: 4-15 <STO>
 R:Mercer, W.D.; Winn, S.I.; Watson, H.C.
 J. Mol. Biol. 104, 277-283, 1976
 A:Title: Twining in crystals of human skeletal muscle D-glyceraledehyde-3-phosphate dehy
 A:Reference number: A38927; MUID:76265083
 A:Accession: A38927
 A:Contents: annotation; X-ray crystallography, 3.5 angstroms
 R:Watson, H.C.; Campbell, J.C.
 submitted to the Brookhaven Protein Data Bank, June 1983
 A:Reference number: A50598; PDB:3GPD
 A:Contents: annotation; X-ray crystallography, 3.5 angstroms, residues 2-8, 'D', 10-40, 'H',
 168-189, 'S', 191-197, 'G', 199-202, 'A', 204-205, 'L', 207-224, 'D', 226-242, 'L', 244-263, 'E', 265-
 C:Comment: Although the gene for this protein is a member of a multigene family and there
 , only one gene appears to be expressed.
 C:Genetics:
 A:Gene: GDB:GAPD
 A:Cross-references: GDB:119249; OMIM:138400
 A:Map position: 12p13.31-12p13.1
 A:Introns: 10/2; 43/3; 79/2; 109/3; 148/2; 175/3; 313/2
 C:Complex: homotrimer; crystallizes as two symmetric dimers in which the partners have
 C:Function: monomer
 C:Function: <CTT>
 A:Description: oxidoreductase; reversibly catalyzes the oxidative phosphorylation of gly
 A:Pathway: gluconeogenesis; glycolysis
 A:Note: tetrameric form; cytosol
 C:Function: <NDC>
 A:Description: DNA repair; uracil DNA glycosylase for base-excision in DNA repair
 A:Note: monomeric form; nuclear
 C:Superfamily: glyceraledehyde-3-phosphate dehydrogenase
 C:Keywords: cytosol; DNA repair; gluconeogenesis; glycolysis; glycosidase; homotrimer;
 F:2-335/Product: glyceraledehyde-3-phosphate dehydrogenase #status experimental <MAT>
 F:5-35/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:152/Active site: Cys #status experimental
 F:173/Active site: His #status predicted

Query Match 88.4%; Score 38; DB 1; Length 335;
 Best Local Similarity 87.5%; Pred. No. 7.3;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 VDGFGRIG 8
 1:|||||
 Db 8 VNGFGRIG 15

Search completed: June 13, 2001, 14:23:13
 Job time: 744 sec